



**Caractérisation et expression de nouveaux éléments  
génétiques transposables de la superfamille Tc1-Mariner  
chez la microalgue marine *Amphora acutiuscula*  
(Bacillariophyta).**

Duc Hung Nguyen

**► To cite this version:**

Duc Hung Nguyen. Caractérisation et expression de nouveaux éléments génétiques transposables de la superfamille Tc1-Mariner chez la microalgue marine *Amphora acutiuscula* (Bacillariophyta).. Biologie végétale. Université du Maine, 2014. Français. NNT : 2014LEMA1013 . tel-01223438

**HAL Id: tel-01223438**

**<https://theses.hal.science/tel-01223438>**

Submitted on 2 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Thèse de Doctorat

**Duc Hung NGUYEN**

*Mémoire présenté en vue de l'obtention du **grade de Docteur de l'Université du Maine** sous le label de L'Université Nantes Angers Le Mans*

École doctorale : **VENAM (Végétal – Environnement – Nutrition – Agroalimentaire – Mer)**  
Discipline : **Biologie**  
Spécialité : **Biologie des Organismes**  
Unité de recherche : **Mer-Molécules-Santé (MMS-2160)**

Soutenue le **17 septembre 2014**  
Thèse N° :

## **Caractérisation et expression de nouveaux éléments génétiques transposables de la superfamille *Tc1-mariner* chez la microalgue marine *Amphora acutiuscula* (Bacillariophyta)**

### **JURY**

Rapporteurs : **Marie-Angèle GRANDBASTIEN**, Directrice de recherche, INRA de Versailles

**Susana COELHO**, Chargée de recherche CNRS, Station biologique de Roscoff

Examineur : **Corinne AUGÉ-GOUILLOU**, Professeur, Université de Tours

Directrice de Thèse : **Annick MORANT-MANCEAU**, Professeur, Université du Mans

Co-directrice de Thèse : **Nathalie CASSE**, MCF-HDR, Université du Mans

Co-encadrante : **Aurore CARUSO**, MCF, Université du Mans

## REMERCIEMENTS

J'ai pu réaliser l'ensemble de mon travail de thèse dans le laboratoire Mer, Molécules, Santé à l'université du Maine (France) grâce à l'obtention d'une bourse octroyée par le Ministère de l'éducation et de la formation du Viêt Nam pendant 3 ans (2010-2013), dans le cadre du programme 322.

Tout d'abord, je tiens à remercier le Professeur Annick MORANT-MANCEAU, Nathalie CASSE et Aurore CARUSO, Maîtres de conférences, pour m'avoir accueilli au sein de votre équipe de recherche et m'avoir encadré durant cette thèse. Je vous suis reconnaissant de m'avoir formé aux techniques de biologie moléculaire et de physiologie végétale, de la confiance que vous m'avez accordée pour mener à bien ce sujet de recherche et d'avoir corrigé mon travail.

Je remercie Mesdames Marie-Angèle GRANDBASTIEN, Susana COELHO et Corinne AUGÉ-GOUILLOU d'avoir accepté de juger ce travail, vos remarques et critiques me seront précieuses.

J'ai bénéficié de l'aide efficace de Dorothee HERMANN pour la construction de l'arbre phylogénétique, je la remercie vivement aussi pour toutes les discussions enrichissantes que nous avons eues ensemble. Je remercie Françoise DENIS, Stanislas THIRIET-RUPERT et Justine MARCHAND qui m'ont permis de compléter mon étude avec l'évaluation du polymorphisme, le nombre de copies et l'expression des éléments transposables. Je tiens à remercier également Jacques-Déric ROUAULT pour son aide concernant la classification des éléments transposables.

Merci à Brigitte MOREAU, Sophie HIARD, Boris JACQUETTE, Emmanuelle TASTARD pour leur aide et leurs conseils techniques. Je voudrais également remercier les autres membres du laboratoire : Benoît CHENAIS, Gérard TREMBLIN, Benoît SCHOEFS, Vincent, Yann, Jean-Luc, Romain pour le partage de connaissances scientifiques et de relations amicales. Je n'oublie pas les doctorants du laboratoire : Fadoumo, Fiddy, Stéphane, Chiraz, Salma, Sahar pour leur amitié.

Merci également aux amis vietnamiens qui font leurs études à l'Université du Maine et aux familles vietnamiennes qui résident au Mans et qui m'ont apporté au quotidien leurs encouragements très amicaux. J'ai toujours eu l'impression de faire partie de ces familles pendant mon séjour en France. Enfin, durant ces trois années de thèse, bien que j'étais loin de ma famille, ma femme NGUYEN Phuong Thao et ma petite fille NGUYEN Phuong Ngoc Diêp m'ont toujours encouragé pour mener ce travail jusqu'à son terme. Je leur suis très reconnaissant.

# TABLE DES MATIERES

<b>LISTE DES FIGURES .....</b>	<b>5</b>
<b>LISTE DES TABLEAUX .....</b>	<b>9</b>
<b>LISTE DES ABREVIATIONS.....</b>	<b>9</b>
<b>INTRODUCTION .....</b>	<b>12</b>
<b>PARTIE I : SYNTHESE BIBLIOGRAPHIQUE .....</b>	<b>14</b>
<b>1. Histoire évolutive des diatomées .....</b>	<b>14</b>
1.1. Évolution des eucaryotes .....	14
1.2. Évolution des eucaryotes photosynthétiques .....	17
1.3. Évolution des diatomées .....	19
<b>2. Généralités sur les diatomées.....</b>	<b>21</b>
2.1. Morphologie et croissance .....	21
2.2. Rôle écologique et valorisation .....	22
2.3. Photosynthèse oxygénique .....	24
2.3.1. <i>Chloroplastes</i> .....	24
2.3.2. <i>Photosynthèse</i> .....	27
2.3.3. <i>Fluorescence de la chlorophylle a</i> .....	29
2.4. Génétique et génome des diatomées.....	31
2.5. Réponses des diatomées aux contraintes de l'environnement.....	32
2.5.1. <i>Définition</i> .....	32
2.5.2. <i>Réponses des diatomées marines aux stress</i> .....	33
<b>3. Éléments transposables .....</b>	<b>36</b>
3.1. Classification .....	36
3.2. Cycle de vie des éléments transposables dans les génomes .....	40
3.3. Impacts des éléments transposables sur les génomes .....	42
3.3.1. <i>Effets délétères des éléments transposables</i> .....	42
3.3.2. <i>Éléments transposables sont-ils bénéfiques ?</i> .....	43
3.4. Utilisation des éléments transposables .....	44
3.4.1. <i>Outils de mutagenèse et de transgénèse</i> .....	44
3.4.2. <i>Marqueurs de polymorphisme</i> .....	45
3.5. Transposons mariner .....	46
3.5.1. <i>MLE chez les animaux</i> .....	46
3.5.2. <i>MLE chez les végétaux</i> .....	47
3.5.3. <i>Éléments transposables chez les microalgues</i> .....	47
3.5.4. <i>Structure des transposons MLE</i> .....	49
3.5.5. <i>Structure protéique des transposases de MLE</i> .....	50
3.5.6. <i>Mécanisme de transposition des MLE</i> .....	52



<b>4. Stress et éléments transposables.....</b>	<b>55</b>
<b>PARTIE II : MATERIELS ET METHODES .....</b>	<b>60</b>
<b>1. Matériel végétal et conditions de culture.....</b>	<b>60</b>
1.1. La diatomée <i>Amphora acutiuscula</i> .....	60
1.2. Milieu de culture.....	61
1.3. Conditions de culture et de stress .....	61
<b>2. Croissance d'<i>Amphora acutiuscula</i> .....</b>	<b>62</b>
2.1. Estimation de la densité cellulaire .....	62
2.2. Courbe de croissance .....	62
2.3. Récolte de la biomasse .....	63
<b>3. Activité photosynthétique .....</b>	<b>63</b>
3.1. Mesures de l'activité photosynthétique par oxymétrie.....	63
3.2. Mesure de la fluorescence de la chlorophylle a.....	65
<b>4. Analyses statistiques .....</b>	<b>68</b>
<b>5. Techniques de biologie moléculaire .....</b>	<b>69</b>
5.1. Extraction des acides nucléiques .....	69
5.2. Contrôle des extractions .....	70
5.3. Amplification par réaction de polymérisation en chaîne (PCR).....	71
5.4. PCR inverse et reconstruction de séquences .....	73
5.5. RT- PCR .....	79
5.6. PCR en temps réel (Q-PCR).....	79
5.7. Southern blot .....	83
5.8. Clonage de fragments d'ADN .....	85
5.9. Séquençage des fragments d'intérêt .....	88
5.10. Bioinformatique.....	88
<b>6. Phylogénie .....</b>	<b>90</b>
6.1. Matériel moléculaire.....	90
6.2. Méthode de phylogénie .....	90
<b>PARTIE III : RESULTATS ET DISCUSSION .....</b>	<b>93</b>
<b>Chapitre 1 : Caractérisation des MLE chez <i>Amphora acutiuscula</i> .....</b>	<b>93</b>
<b>1. Détection de séquences partielles de MLE dans le génome de <i>A. acutiuscula</i> .....</b>	<b>93</b>

<b>2. Allongement des séquences de MLE par PCR inverse .....</b>	<b>97</b>
2.1. Séquences obtenues par PCR inverse.....	97
2.2. Allongement de la séquence Re-1 en régions 5' et 3' .....	104
2.3. Allongement des séquences Re-2 vers l'extrémité 5' .....	115
<b>3. Caractérisation du transposon <i>Aacutmar1</i> dans le génome d'<i>A. acutiuscula</i> .....</b>	<b>119</b>
3.1. Caractérisation des ITR d' <i>Aacutmar1</i> .....	119
3.2. Caractérisation de séquences complètes des transposons <i>Aacutmar1</i> .....	121
3.2.1. Identification des MLE <i>Aacutmar1</i> complets à l'aide des amorces définies sur les ITR.....	121
3.2.2. Établissement d'un consensus pour les éléments <i>Aacutmar1</i> .....	124
3.2.3. Caractérisation des séquences complètes <i>Aacutmar1</i> .....	126
3.2.4. Comparaison de <i>ConsAacutmar1</i> avec des MLE de plantes disponibles dans les banques de données .....	128
3.3. Estimation du nombre de copies d' <i>Aacutmar</i> dans le génome de <i>A. acutiuscula</i> ....	132
3.3.1. Visualisation des copies d' <i>Aacutmar</i> par Southern blot .....	132
3.3.2 Estimation du nombre de copies par PCR quantitative .....	134
3.4. Relations phylogénétiques des MLE d' <i>A. acutiuscula</i> avec ceux des plantes et des autres Straménopiles présents dans les banques de données.....	138
<b>4. Allongement de l'extrémité 3' de l'élément Re-3.2 .....</b>	<b>141</b>
<b>5. Discussion .....</b>	<b>147</b>
5.1. Présence des MLE dans le génome d' <i>A. acutiuscula</i> et obtention d'un élément complet .....	147
5.2. Amplification des MLE dans le génome d' <i>A. acutiuscula</i> avec les amorces d'ITR	148
5.3. Comparaison des séquences nucléotidiques d' <i>Aacutmar1</i> à celles de MLE végétaux des banques de données .....	148
5.3.1. Le TSD et les ITR des éléments <i>Aacutmar1</i> .....	149
5.3.2. Structure de la transposase <i>AaTnp</i> .....	151
5.4. Dénombrement des copies d' <i>Aacutmar</i> dans le génome hôte .....	152
5.5. Relations phylogénétiques d' <i>Aacutmar1</i> .....	155
5.5.1. Révision de la nomenclature du groupe formé par les MLE de végétaux et de Straménopiles .....	155
5.5.2. Relations phylogénétiques entre les MLE des diatomées et les MLE dans la lignée verte .....	156
<b>Chapitre 2 : Impact de variations de température et d'excès de métaux sur l'expression des MLE .....</b>	<b>160</b>
<b>1. Impact des traitements thermiques .....</b>	<b>160</b>
1.1. Activité photosynthétique.....	160
1.1.1. Intensité photosynthétique .....	160
1.1.2. Fluorescence de la chlorophylle <i>a</i> .....	161
1.2. Détection de l'expression des MLE par RT-PCR .....	162

1.2.1. <i>Extraction des ARN totaux</i> .....	162
1.2.2. <i>RT-PCR</i> .....	164
1.3. Détection par Q-RT-PCR de l'expression des MLE et des gènes de réponse aux stress .....	170
1.3.1. <i>Caractérisation des gènes de ménage</i> .....	170
1.3.2. <i>Caractérisation des gènes d'intérêt</i> .....	170
1.3.3. <i>PCR en temps réel</i> .....	173
<b>2. Impact d'un excès de zinc</b> .....	<b>179</b>
2.1. <i>Activité photosynthétique</i> .....	179
2.2. <i>Détection de l'expression des MLE</i> .....	181
<b>3. Impact d'un excès de cuivre</b> .....	<b>184</b>
3.1. <i>Activité photosynthétique</i> .....	184
3.2. <i>Détection de l'expression des MLE</i> .....	185
<b>4. Discussion</b> .....	<b>187</b>
4.1. <i>Effets des variations de température</i> .....	187
4.2. <i>Effet d'un excès de zinc</i> .....	193
<b>CONCLUSIONS ET PERSPECTIVES</b> .....	<b>195</b>
<b>BIBLIOGRAPHIE</b> .....	<b>197</b>
<b>ANNEXES</b> .....	<b>230</b>

## LISTE DES FIGURES

Figure 1 : Arbre phylogénétique des principaux groupes eucaryotes établi par Baldauf (2008) d'après des données moléculaires et ultrastructurales. ....	14
Figure 2 : Phylogénie non enracinée des trois mégagroupes chez les eucaryotes d'après Burki et al. (2008).....	16
Figure 3 : Représentation de l'origine des plastes chez les diatomées grâce aux endosymbioses primaire (A) et secondaire (B) et leurs effets potentiels sur l'évolution du génome.D'après Armbrust (2009).....	17
Figure 4 : Frustule d'une diatomée du genre <i>Achnantes</i> (Microscopie électronique à balayage - CRPGL, Luxembourg).....	21
Figure 5 : A-Coupe dans des cellules de <i>Phaeodactylum tricornutum</i> renfermant deux chloroplastes observés en microscopie électronique (d'après Falkowski et Raven, 1997). B-Schéma de la structure d'un chloroplaste de diatomée avec des thylakoïdes empilés par trois. ....	25
Figure 6 : Schéma des relations entre les principaux groupes d'organismes photosynthétiques oxygénés, les étapes évolutives possibles et leurs types d'antennes collectrices de la lumière.D'après Niyogi et Truong(2013). ....	26
Figure 7 : Schéma de l'organisation de la membrane d'un thylakoïde. ....	28
Figure 8 : Schéma des voies de dissipation de l'énergie lumineuse au niveau du photosystème II. ....	30
Figure 9 : A- Sites de production des formes réactives de l'oxygène (ROI) dans une cellule végétale ; B- Production d'ion superoxyde ( $O_2^-$ ) dans le chloroplaste d'une plante et voie de détoxification enzymatique. SOD : superoxyde dismutase ; APX : ascorbate peroxydases ; CAT : catalase. D'après Mitler (2002).....	35
Figure 10 : Schéma de la classification des ET selon leur mode de transposition. ....	37
Figure 11 : Classification des éléments transposables (d'après Wicker et al., 2007).....	38
Figure 12 : Évolution du cycle de vie des éléments <i>Tc1/mariner</i> chez les hôtes naturels. D'aprèsMiskey et al. (2005).....	42
Figure 13 : Structure des éléments <i>mariner</i> . ....	49
Figure 14 : Mécanisme de transposition de <i>mariner</i> . D'après Jaillet et al. (2012). ....	53
Figure 15 : Induction du déplacement des éléments transposables après un stress(D'après Capy et al., 2000).....	55
Figure 16 : Impact des facteurs de l'environnement sur les éléments transposables et le génome (Biémont et Vieira, 2006). ....	57
Figure 17 : Localisation de Càn Gio (latitude: 10°40'09'' ; longitude: 107°00'59'') sur la côte Sud-est du Viêtnam.....	60
Figure 18 : Photographie de microscopie électronique à balayage de la microalgue marine <i>Amphora acutiuscula</i> (cliché de P. Gaudin et T.L.N. Nguyen-Deroche). ....	60
Figure 19 : Courbe de croissance théorique des diatomées cultivées en batch en fonction du temps.....	62
Figure 20 : Vue éclatée de la cellule de mesure de l'oxymètre.....	64
Figure 21 : Vue de dessus de l'extension du fluorimètre modifié par Rech et al. (2003). ....	66
Figure 22 : Évolution de la fluorescence chlorophyllienne d'une suspension d'A. <i>acutiuscula</i> à 24 °C.....	67
Figure 23 : Allongement d'une séquence connue par PCR inverse. ....	74
Figure 24 : Schéma récapitulatif des étapes de la reconstruction des séquences MLE chez A. <i>acutiuscula</i> . ....	76
Figure 25 : Schéma de l'emplacement des amorces permettant de vérifier la présence réelle des séquences reconstruites dans le génome d'A. <i>acutiuscula</i> . ....	78
Figure 26 : Schéma du principe du transfert d'ADN sur membrane par Southern blot. ....	84

Figure 27 : Schéma du vecteur pGEM®-T Easy.....	86
Figure 28 : Schéma de la démarche informatique simplifiée pour détecter les ITR. ....	89
Figure 29 : Schéma général d'un MLE isolé à partir d'un ADNg végétal, sur lequel sont localisées les amorces <i>MLE5A</i> et <i>MLE3A</i> ayant servi à la détection des séquences chez <i>A. acutiuscula</i> . ....	93
Figure 30 : Électrophorèse sur gel d'agarose 1 % des produits d'amplification obtenus avec les amorces dégénérées <i>MLE5A</i> et <i>MLE3A</i> chez la diatomée <i>A. acutiuscula</i> . ....	94
Figure 31 : Alignement des séquences partielles <i>Aacutmar-sp1.1</i> à <i>1.8</i> traduites en protéines. ....	96
Figure 32 : Schéma permettant de localiser les amorces ayant servi à la recherche des MLE complets chez la diatomée <i>A. acutiuscula</i> . ....	100
Figure 33 : Migration sur gel d'agarose des produits d'amplification pour l'obtention des fragments flanquant la séquence <i>ConsAacutmar-sp1</i> . ....	101
Figure 34 : A) Schéma de l'alignement de la séquence <i>ConsAacutmar-sp1</i> avec les fragments <i>EcoRIR2F4-5'</i> et <i>EcoRIR2F4-3'</i> . B) L'assemblage des trois fragments, dans l'ordre, <i>EcoRIR2F4-5'</i> , <i>ConsAacutmar-sp1</i> et <i>EcoRIR2F4-3'</i> a permis de reconstruire la séquence nommée <i>Re-1</i> . ....	102
Figure 35 : Traduction de la séquence <i>Re-1</i> en acides aminés. ....	103
Figure 36 : Schéma permettant de localiser les amorces ayant servi à l'allongement du fragment reconstruit <i>Re-1</i> de 1268 pb. ....	104
Figure 37 : Migration sur gel d'agarose des produits d'amplification pour l'obtention des fragments flanquant la séquence <i>Re-1</i> . ....	104
Figure 38 : A) Schéma de l'alignement de la séquence <i>Re-1</i> avec les fragments <i>ApaIR5F6-5'</i> et <i>ApaIR5F6-3'</i> . B) L'assemblage de <i>ApaIR5F6-5'</i> et <i>ApaIR5F6-3'</i> avec <i>Re-1</i> , d'une part et de <i>ApaIR5F6-5'</i> avec <i>Re-1</i> d'autre part a permis de reconstruire deux séquences différentes. ....	106
Figure 39 : Alignement des séquences reconstruites <i>Re-2.1</i> et <i>Re-2.2</i> . ....	108
Figure 40 : Traduction en protéine de l'extrémité 5' identique des deux séquences <i>Re-2.1</i> et <i>-2.2</i> et positionnement des motifs caractéristiques des transposases. ....	109
Figure 41 : Schéma de la localisation des amorces utilisées pour vérifier la présence des séquences des éléments reconstruits <i>Re-2.1</i> et <i>Re-2.2</i> . ....	110
Figure 42 : Amplifications à partir de l'ADNg d' <i>A. acutiuscula</i> des fragments de vérification des régions 3' des séquences <i>Re-2.1</i> et <i>-2.2</i> . ....	111
Figure 43 : Alignement du fragment <i>F6int/VRe-2.1</i> avec la région 3' du fragment reconstruit <i>Re-2.1</i> . ....	112
Figure 44 : Alignement du fragment de vérification <i>F6int/VRe-2.2</i> avec la région 3' de la séquence reconstruite <i>Re-2.2</i> . ....	113
Figure 45 : Schéma permettant de localiser les amorces ayant servi à l'allongement de l'extrémité 5' de <i>Re-2.1</i> (A) et <i>Re-2.2</i> (B). ....	114
Figure 46 : Amplification par iPCR du fragment bordant l'extrémité 5' des séquences <i>Re-2</i> . ....	114
Figure 47 : Comparaison des séquences des ITR-5' et -3' détectées dans l'élément reconstruit <i>Re-3.1</i> d' <i>A. acutiuscula</i> . ....	116
Figure 48 : A) Alignement du fragment <i>EcoRVR10F13-5'</i> avec les deux séquences <i>Re-2.1</i> et <i>Re-2.2</i> . B) L'assemblage du fragment <i>EcoRVR10F13-5'</i> avec <i>Re-2.1</i> et <i>-2.2</i> a permis de former les séquences <i>Re-3.1</i> et <i>Re-3.2</i> , respectivement. ....	117
Figure 49 : Séquences nucléique et protéique du transposon <i>Aacutmar1-recomplet</i> identifié à partir de la séquence reconstruite <i>Re-3.1</i> . ....	118
Figure 50 : Schéma de l'élément <i>Re-3.1</i> montrant la localisation des amorces définies pour vérifier les ITR de l'élément <i>Aacutmar1</i> dans le génome d' <i>A. acutiuscula</i> . ....	119

Figure 51 : Migration sur gel d'agarose des produits de PCR contenant les ITR de l'élément <i>Aacutmar1</i> dans le génome d' <i>A. acutiuscula</i> .	120
Figure 52 : Alignement des séquences des ITR-5' (A) et -3' (B) obtenues par iPCR dans l'élément reconstruit <i>Re-3.1</i> avec ceux obtenus lors de la vérification par PCR à partir de l'ADNg d' <i>A. acutiuscula</i> .	120
Figure 53 : Comparaison des ITR-5' et -3' d' <i>A. acutiuscula</i> avec ceux des éléments complets disponibles dans la littérature et dans les banques de données.	121
Figure 54 : Schéma de l'amplification des éléments complets <i>Aacutmar1</i> avec les amorces spécifiques et dégénérées définies sur les ITR.	122
Figure 55 : Migration sur gel d'agarose des produits de PCR obtenus avec les amorces établies sur les ITR d' <i>Aacutmar1-re</i> .	122
Figure 56 : Structure et organisation des différents éléments <i>Aacutmar</i> présents dans le génome d' <i>A. acutiuscula</i> .	124
Figure 57 : Séquence consensus du transposon <i>ConsAacutmar1</i> .	127
Figure 58 : Alignement d' <i>AaTnp</i> avec les transposases d'autres éléments complets de la famille <i>mariner</i> .	130
Figure 59 : Alignement des motifs NLS des transposases de plantes et de diatomées.	131
Figure 60 : Migration sur gel d'agarose d'agarose 1 % des ADNg d' <i>A. acutiuscula</i> .	132
Figure 61 : Schéma de la localisation de la sonde de 400 pb et des deux séquences témoins.	133
Figure 62 : Southern blot obtenu à partir d'un transfert d'ADNg et de témoins d'hybridation.	133
Figure 63: Courbe standard obtenue avec les amorces <i>acutimarF1/R1</i> sur une gamme de concentrations en plasmides contenant un fragment du gène <i>Aacutmar1</i> .	135
Figure 64 : Courbe standard obtenue sur ADNg avec les amorces ciblant le gène de la transposase d' <i>Aacutmar1</i> .	135
Figure 65 : Relations phylogénétiques entre les transposases complètes des MLE de Straménopiles et les MLE des plantes angiospermes terrestres.	139
Figure 66 : Schéma permettant de localiser les amorces ayant servi à l'allongement de l'extrémité 3' de l'élément <i>Re-3.2</i> .	141
Figure 67 : Migration sur gel d'agarose des produits d'amplification du fragment bordant l'extrémité 3' de la séquence <i>Re-3.2</i> .	142
Figure 68 : A) Schéma de l'élément reconstruit <i>Re-4</i> montrant la localisation des amorces définies pour vérifier sa présence dans le génome d' <i>A. acutiuscula</i> . B) Amplification du fragment de vérification de la présence de <i>Re-4</i> .	143
Figure 69 : A) Schéma de l'alignement du fragment <i>NspIR6F8-3'</i> avec la séquence <i>Re-3.2</i> . B) L'assemblage du fragment <i>NspIR6F8-3'</i> avec <i>Re-3.2</i> a permis de former la séquence <i>Re-4</i> .	144
Figure 70 : Alignement de <i>ConsAacutmar1</i> et d' <i>Aacutmar4</i> .	146
Figure 71 : Schéma récapitulatif des deux hypothèses pouvant expliquer la transmission des MLE entre les algues vertes et les Straménopiles (d'après la figure de Tirichine et Bowler, 2011).	157
Figure 72 : Électrophorèse sur gel d'agarose 1 % des produits d'extraction d'ARN totaux d' <i>A. acutiuscula</i> placée pendant 0 à 8 h à 16 ou 32 °C. Sont présentés ici 3 réplicats de la cinétique : T= Température témoin 24 °C.	163
Figure 73 : Schéma montrant la position des trois couples d'amorces ayant servi à la détection de l'expression des MLE dans les différentes conditions thermiques.	164
Figure 74 : Amplification des gènes de contrôle, de réponse aux stress de 16 °C et des MLE par RT-PCR.(Exemple de résultats pour le stress 16 °C, 1 <sup>er</sup> réplicat).	165
Figure 75 : Amplification des gènes de ménage, de réponse aux stress de 32 °C et des MLE par RT-PCR.(Exemple de résultats obtenus à 32 °C, 1 <sup>er</sup> réplicat)	166

Figure 76 : Alignement en acides nucléiques des 5 séquences ADNc de MLE ( <i>F9-Stop-1</i> à -5) et des 6 séquences pleine taille et à cadre de lecture ouvert ( <i>Aacutmar1</i> ) isolées dans le présent travail et présentées dans le chapitre précédent. ....	169
Figure 77 : Électrophorèse sur gel d'agarose 1 % des produits d'amplification obtenus avec les amorces <i>QHSPampfw</i> / <i>QHSPAmprw</i> chez <i>A. acutiuscula</i> .....	171
Figure 78 : Alignement nucléotidique des six séquences obtenues avec les trois couples d'amorces ayant servi à la Q-RT-PCR et ciblant le gène de la transposase des MLE avec les deux séquences des banques de données les plus proches <i>Aacutimal.4</i> et <i>1.6</i> . ....	172
Figure 79 : Alignement protéique des six séquences obtenues avec les trois couples d'amorces utilisées en Q-RT-PCR et ciblant le gène de la transposase des MLE avec les deux séquences déposées dans les banques de données et obtenues dans thèse D. Hermann (2011).....	172
Figure 80 : Courbe de dissociation obtenue avec le premier couple d'amorces ciblant le gène de l'ubiquitine. La présence de deux pics traduit celle de deux amplicons.....	174
Figure 81 : Courbe standard obtenue avec le deuxième couple d'amorces ciblant l'ubiquitine. ....	174
Figure 82 : Courbe de dissociation obtenue avec le couple d'amorces ciblant le gène 18S.....	175
Figure 83 : Expression relative du gène de la hsp70 chez <i>A. acutiuscula</i> placée pendant 0,5 à 8 h à 16, 24 ou 32 °C comparativement au témoin (24 °C).....	177
Figure 84 : Expression relative du gène de la transposase des MLE chez <i>A. acutiuscula</i> placée pendant 0,5 à 8 h à 16, 24 ou 32 °C comparativement au témoin (24 °C).....	178
Figure 85 : A-Évolution de l'intensité photosynthétique. B-Évolution du rendement quantique maximal (Fv/Fm) et du rendement quantique effectif (ΦPSII) du photosystème II chez <i>A. acutiuscula</i> en présence d'un excès de zinc (20 μM) pendant 0 à 8 h.....	179
Figure 86 : Évolution des rendements quantiques maximaux (Fv/Fm) et effectifs (ΦPSII) du photosystème II chez <i>A. acutiuscula</i> placée pendant 8 h dans de l'eau de mer enrichie en zinc.....	180
Figure 87 : Électrophorèse des ARN totaux extraits d' <i>A. acutiuscula</i> exposée pendant 0,5 à 8 h à un excès de zinc (40 μM). 1 <sup>er</sup> et 2 <sup>ème</sup> répliquats.....	181
Figure 88 : Expression relative du gène de la hsp70 chez <i>A. acutiuscula</i> exposée pendant 0,5 à 8 h à un excès de zinc (40 μM) dans le milieu de culture comparativement au témoin. ....	182
Figure 89 : Expression relative du gène de la transposase des MLE chez <i>A. acutiuscula</i> exposée pendant 0,5 à 8 h à un excès de zinc (40 μM) dans le milieu de culture comparativement au témoin. ....	183
Figure 90 : A : Évolution de l'intensité photosynthétique. B : Évolution du rendement quantiques maximal (Fv/Fm) et du rendement quantique effectif (ΦPSII) du photosystème II chez <i>A. acutiuscula</i> en présence de cuivre 3 μM pendant 0 à 8 heures. Expérience réalisée une seule fois.....	184
Figure 91 : Électrophorèse des ARN totaux extraits d' <i>A. acutiuscula</i> exposée à un excès de cuivre (3 μM) pendant 0,5 à 8 h (exemple d'un répliquat).....	185

## LISTE DES TABLEAUX

Tableau 1 : Tableau récapitulatif des couples d'amorces dégénérées ou spécifiques des différents gènes étudiés. ....	72
Tableau 2 : Tableau récapitulatif des enzymes et amorces utilisées en iPCR. ....	75
Tableau 3 : Tableau récapitulatif des couples d'amorces utilisés pour vérifier des fragments reconstruits et amplifier les séquences contenant ITR, les MLE complets. ....	77
Tableau 4 : Tableau récapitulatif des gènes étudiés et leurs amorces pour la Q-PCR. ....	81
Tableau 5 : Tableau récapitulatif des ET complets de la superfamille <i>Tc1-mariner</i> et <i>ITm</i> utilisés pour l'étude phylogénétique. ....	91
Tableau 6 : Tableau récapitulant les différentes séquences partielles de MLE obtenues chez la diatomée <i>A. acutiuscula</i> . ....	95
Tableau 7 : Comparaison deux à deux des huit séquences d' <i>Aacutmar-sp1</i> disponibles au laboratoire. ....	95
Tableau 8 : Tableau récapitulant les caractéristiques des 12 séquences complètes des éléments <i>Aacutmar1</i> et de leur consensus obtenus chez la diatomée <i>A. acutiuscula</i> . ....	125
Tableau 9 : Comparaison deux à deux des 12 séquences complètes d' <i>Aacutmar1</i> . ....	125
Tableau 10 : Liste des transposases putatives dont la séquence protéique a été comparée avec <i>ConsAacutmar</i> et leur pourcentage d'homologie avec <i>ConsAacutmar</i> . ....	129
Tableau 11 : Séquences des motifs Hélice-Tour-Hélice de transposases de plantes et de diatomées. ....	129
Tableau 12 : Identification des motifs NLS prédits dans les transposases de plantes et de diatomées. ....	131
Tableau 13 : Résultats de l'estimation du nombre de copies de MLE dans le génome d' <i>A. acutiuscula</i> NNC 216 par PCR en temps réel. ....	137
Tableau 14 : La longueur des ITR, des UTR, de l'ORF des MLE complets chez les végétaux. ....	149
Tableau 15 : Effet de la température sur l'intensité photosynthétique ( $\mu\text{mol O}_2 \text{ h}^{-1} \text{ mg}^{-1}$ Chl <i>a</i> ) de la diatomée <i>A. acutiuscula</i> maintenue pendant 0 à 8 h à 16, 24 et 32 °C. ....	160
Tableau 16 : Effet de la température sur le rendement quantique maximum (Fv/Fm) chez <i>A. acutiuscula</i> maintenue pendant 0 à 8 h à 16, 24 et 32 °C. ....	161
Tableau 17 : Effet de la température sur le rendement quantique effectif ( $\Phi\text{PSII}$ ) chez <i>A. acutiuscula</i> maintenue pendant 0 à 8 h à 16, 24 et 32 °C. ....	161
Tableau 18 : Liste des couples d'amorces utilisés en PCR en temps réel et conditions retenues. ....	175
Table 19 : Récapitulatif des réponses significativement différentes de celles des cultures témoins observées chez <i>A. acutiuscula</i> exposée à différentes températures ou à un excès de zinc pendant 0,5 à 8 h. ....	186



## LISTE DES ABREVIATIONS

ADN : acide désoxyribonucléique  
ADNc : acide désoxyribonucléique complémentaire  
ADNg : acide désoxyribonucléique génomique  
AP : activité photosynthétique  
ARNm : acide ribonucléique messenger  
BET : bromure d'éthidium  
Chl *a* : chlorophylle *a*  
Chl *b* : chlorophylle *b*  
Chl *c* : chlorophylle *c*  
CTAB : cetyl trimethylammonium bromide  
DEPC : diéthyl pyrocarbonate  
DO : densité optique  
EGT : transfert de gènes endosymbiotique  
EMA : eau de mer artificielle  
ERO : espèces réactives de l'oxygène  
ET : élément transposable  
ER : expression relative  
Fo : fluorescence minimale  
Fm : fluorescence maximale  
Fm' : fluorescence maximale à la lumière  
Fv : fluorescence variable  
Fv/Fm : rendement quantique maximal du PSII  
Fs : Fluorescence de base  
HSP : heat shock proteins  
IP : intensité photosynthétique  
ITR: inverted terminal repeats  
MLE: *mariner*-like element  
NCC : Nantes culture collection  
NPQ : non-photochemical quenching  
ORF: open reading frame  
PAE: Percentile amplification efficiency  
PSRF: (Potential Scale Reduction Factor)  
pb : paire de bases  
ΦPSII : rendement quantique effectif du PSII  
PCR : réaction de polymérisation en chaîne  
PSII : photosystèmes II  
Q-PCR : réaction de polymérisation en chaîne quantitative  
RT-PCR: reverse transcription PCR  
sHSP: small heat shock proteins  
SOD: superoxyde dismutase  
UTR: untranslated region  
UV : ultraviolet

## LISTE DES ANNEXES

Annexe 1 : Eau de mer artificielle (EMA) .....	230
Annexe 2 : Axénie des cultures d' <i>Amphora acutiuscula</i> .....	231
Annexe 3 : Alignement des 12 séquences nucléotidiques d' <i>Aacutmar1</i> . ....	232
Annexe 4 : Alignement de 12 ORF d' <i>Aacutmar1</i> . ....	236
Annexe 5 : Alignement de <i>Re-4</i> et la séquence de sa vérification <i>BATG/VRe-4</i> . ....	237

## INTRODUCTION

Les éléments génétiques transposables (ET) sont des séquences d'ADN, parfois capables de se déplacer dans les génomes et d'induire des mutations. C'est ainsi qu'ils ont été découverts chez le maïs par B. McClintock à la fin des années 1940 comme responsables de modifications instables de la couleur des grains (McClintock, 1953). Des ET ont ensuite été détectés dans tous les génomes dans lesquels ils ont été recherchés ; ils représentent une part plus ou moins importante de ces génomes, de quelques pour cent chez des microalgues marines à paroi siliceuse (diatomées) (Armbrust et al., 2004 ; Bowler et al., 2008) à près de 80 % dans le genre *Lilium* (Capy, 2004 ; Feschotte et al., 2002). Les ET se différencient par des structures et des mécanismes de transposition différents (Capy, 2004 ; Finnegan, 1989 ; Kapitonov et Jurka, 2008 ; Wickler et al., 2007). La plupart des ET sont inactifs mais certains sont capables de se déplacer (transposer) dans le génome dans certaines conditions, en particulier des conditions de stress (Grandbastien, 1998 ; Hashida et al., 2003 ; Maumus et al., 2009 ; Mhiri et al., 1997). Ainsi, les ET jouent un rôle dans la structuration et le fonctionnement des génomes. Ils engendrent une variabilité génétique qui aurait permis à des populations, notamment de diatomées (Maumus et al., 2009 ; Tirichine et Bowler, 2011 ; Vardi et al., 2008), de s'adapter à des changements environnementaux (Feschotte et Pritham, 2007 ; Zhou et al., 2011).

Parmi les producteurs primaires de la planète, les microalgues marines occupent une place prépondérante dans les réseaux trophiques marins. Les diatomées qui représentent environ 40 % du phytoplancton marin, constituent l'un des taxons (groupe des Bacillariophyta) les plus diversifiés et les plus abondants (Round et al., 1990 ; Simon et al., 2009 ; Wilhelm et al., 2006). Ces microalgues sont soumises à des variations incessantes des facteurs de leur environnement comme le niveau d'éclairement, la qualité ou la température de l'eau.

Dans un travail précédent (Hermann, 2011) réalisé au laboratoire Mer, Molécules, Santé à l'université du Maine, des fragments de 400 pb d'ET de la famille *Tc1-mariner* ont été amplifiés par PCR chez plusieurs diatomées marines dont *Amphora acutiuscula*. De plus, l'expression d'éléments *mariner-like* (MLE) a été mise en évidence lors de stress thermiques chez trois espèces de diatomées, *Phaeodactylum tricornutum*, *Amphora coffeaeformis* et *Amphora acutiuscula*, suggérant la présence de MLE potentiellement complet(s) et actif(s).

En se focalisant sur l'espèce *A. acutiuscula* qui s'est révélée être un modèle facile à utiliser pour les techniques de biologie moléculaire, les objectifs du présent travail ont été de poursuivre les travaux de Hermann (2011) afin :

- d'obtenir un ou plusieurs éléments MLE complets.
- caractériser ces éléments et les classer par rapport aux autres MLE, notamment ceux présents chez les plantes.
- préciser la cinétique d'induction des MLE dans des conditions de chocs thermiques mais aussi de stress métalliques. En effet, il a été montré au laboratoire que des concentrations élevées en cuivre et en zinc affectent les paramètres physiologiques d'*A. acutiuscula* (Nguyen-Deroche et al., 2009 ; Nguyen-Deroche et al., 2012).

Pour atteindre ces objectifs, nous avons utilisé la technique de PCR inverse pour allonger les fragments de 400 pb qui avaient déjà été identifiés et des techniques de bioinformatique pour rechercher des motifs protéiques et classer les MLE obtenus. L'impact des chocs thermiques et métalliques sur cette diatomée a été évalué par des mesures d'activité photosynthétique, l'expression des gènes et le nombre de copies de MLE ont été quantifiés par PCR quantitative.

# PARTIE I : SYNTHÈSE BIBLIOGRAPHIQUE

## 1. Histoire évolutive des diatomées

### 1.1. Évolution des eucaryotes

La classification des êtres vivants eucaryotes a longtemps reposée sur des critères morphologiques, biologiques et biochimiques. Les protistes (organismes unicellulaires) étaient alors séparés des organismes pluricellulaires mais, depuis les années 1990, le développement de la systématique moléculaire a engendré des confirmations mais aussi de profonds remaniements des classifications plus anciennes. La construction d'un arbre phylogénétique solide indiquant les relations entre tous les eucaryotes, pluricellulaires et unicellulaires, n'est pas complètement finalisée car des données moléculaires sont absentes ou insuffisantes pour certains organismes et groupes, notamment les unicellulaires (Burki et al, 2007), certains organismes sont mal identifiés et la biodiversité n'est pas encore complètement connue. Les gènes utilisés pour établir les relations phylogénétiques sont nombreux, par exemple : 6-phosphogluconate déshydrogénase, 18S, ITS ADNr, cytochrome oxydase 1 (cox1) et le gène de la rbcL qui code pour la plus grande sous-unité de la ribulose biphosphate carboxylase oxygénase (RuBisCO, enzyme de la photosynthèse) plus spécifiquement pour les organismes photosynthétiques (Evans et al., 2007).

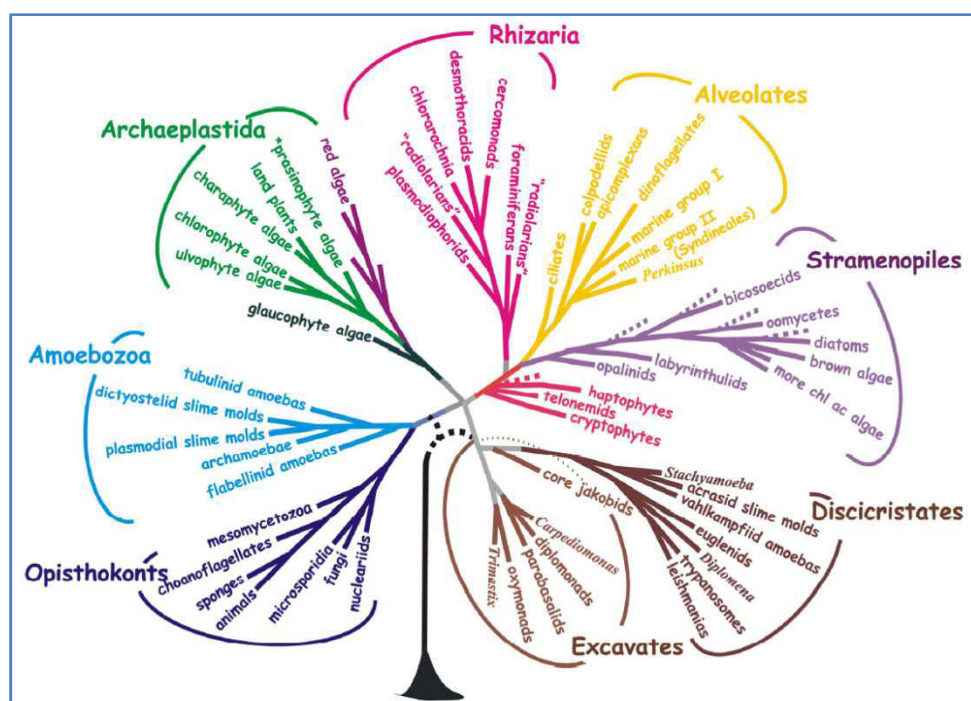


Figure 1 : Arbre phylogénétique des principaux groupes eucaryotes établi par Baldauf (2008) d'après des données moléculaires et ultrastructurales.

En 2006, Parfrey et al. ont montré que la division du monde vivant en 6 règnes ou supergroupes présentait des imperfections. Baldauf en 2008 publie une synthèse bibliographique et un arbre schématisant les relations entre les grands groupes eucaryotes (Fig. 1). Les relations entre les microorganismes eucaryotes ont longtemps été un sujet de controverses. Hackett et al. (2007) ont confirmé la monophylie du groupe des Chromalvéolates en utilisant plusieurs gènes nucléaires exprimés. Ce groupe rassemble des microalgues mais aussi des microorganismes non photosynthétiques. Récemment, l'analyse phylogénomique de Burki et al. (2008) effectuée sur 65 espèces et avec 135 gènes montre que les eucaryotes peuvent être divisés en deux mégagroupes monophylétiques et des lignées moins diversifiées correspondant aux Excavates. Selon ces auteurs, les eucaryotes seraient distribués en 3 mégagroupes monophylétiques (Fig. 2) : i) Excavates, ii) Unikonts (Amoebozoa, champignons et animaux), iii) plantes, stramenopiles, Alveolates, Rhizaria, Haptophytes et Cryptophytes. Ce dernier mégagroupe rassemble presque tous les organismes photosynthétiques à l'exception notable des Euglenozoa rangés dans le groupe des Excavates.

Notre modèle d'étude, *A. acutiuscula*, est une microalgue appartenant au groupe des Straménopiles (ou Hétérokontes)). Par conséquent, notre étude bibliographique se focalisera sur ce groupe complexe.

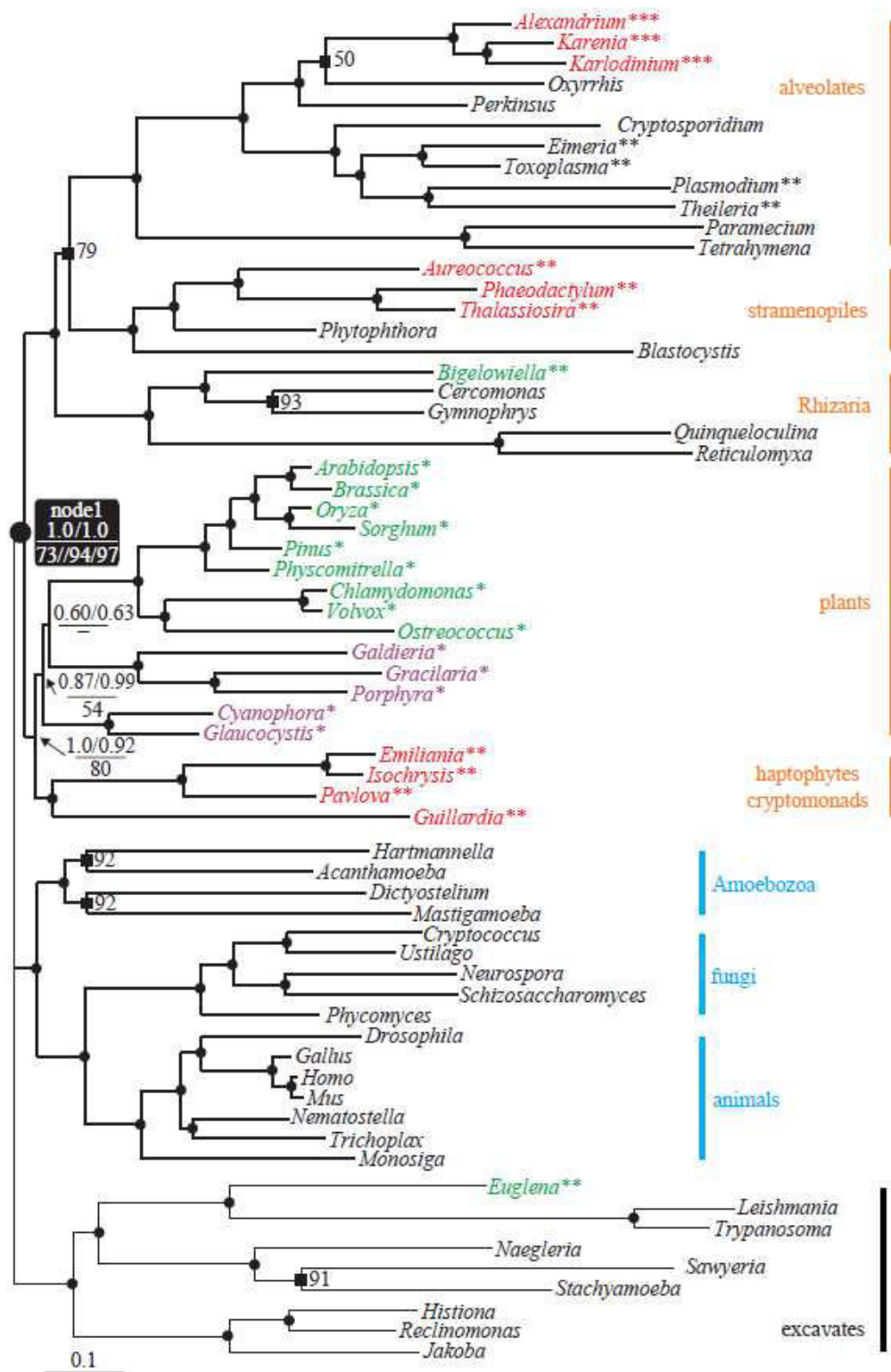
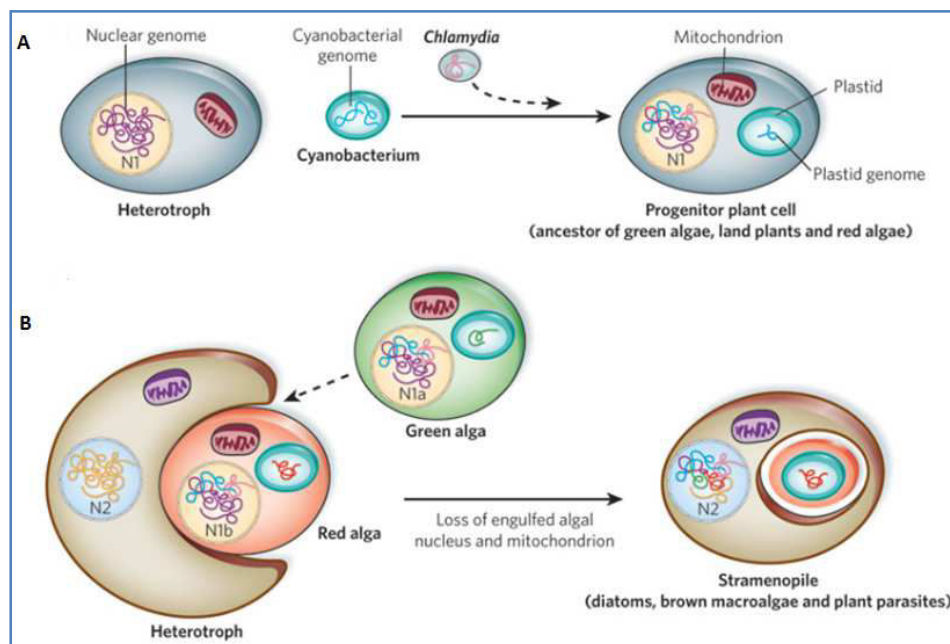


Figure 2 : Phylogénie non enracinée des trois mégagroupes chez les eucaryotes d'après Burki et al. (2008).

Le code de couleur des espèces correspond au type de pigments plastidiaux comme suit: violet: chlorophylle *a* ; vert: chlorophylle *a* + *b* ; rouge : chlorophylle *a* + *c*. Les astérisques représentent les endosymbioses primaire, secondaire ou tertiaire. L'échelle en bas à gauche représente une estimation du nombre de substitutions d'acides aminés par site.

## 1.2. Évolution des eucaryotes photosynthétiques

Les eucaryotes photosynthétiques forment un ensemble d'organismes variés, uni- ou pluricellulaires, vivant dans des milieux très diversifiés mais qui renferment tous de la chlorophylle *a* (Chl *a*) qui est le pigment photosynthétique indispensable pour convertir l'énergie lumineuse en énergie chimique. Chez les eucaryotes photosynthétiques, la Chl *a* localisée dans les chloroplastes est associée à d'autres molécules de Chl (*b*, *c* ou *d*) et à des pigments photorécepteurs de type caroténoïdes ou des phycobiliprotéines (algues rouges). Il est maintenant admis que le chloroplaste de l'ancêtre des plantes, des algues vertes et des algues rouges résulterait d'une endosymbiose primaire (Fig. 3A) au cours de laquelle une cyanobactérie aurait été phagocytée par une cellule hôte eucaryote déjà pourvue de mitochondries (Falkowski et al., 2004 ; Keeling et al., 2004 ; Reyes-Prieto et al., 2007). Les premiers eucaryotes photosynthétiques seraient apparus au Protérozoïque (Précambrien), il y a 1,5 milliards d'années (Falkowski et al., 2004).



**Figure 3 : Représentation de l'origine des plastides chez les diatomées grâce aux endosymbioses primaire (A) et secondaire (B) et leurs effets potentiels sur l'évolution du génome. D'après Armbrust (2009).**

**A :** Lors de l'endosymbiose primaire, une grande partie du génome de la cyanobactérie a été transférée vers le noyau de la cellule hôte (N1), quelques gènes sont restés à l'intérieur du plastide. L'invasion potentielle de l'hôte par un parasite *Chlamydia* est indiquée par une flèche en pointillés, et le transfert de gènes *Chlamydia* vers le noyau de l'hôte est indiqué en rose. Cette cellule ancestrale a ensuite évolué et serait à l'origine des plantes, des algues rouges et des algues vertes qui ont des génomes plastidiaux similaires. **B :** Au cours de l'endosymbiose secondaire, un autre hétérotrophe a phagocyté une algue rouge. La phagocytose potentielle d'une microalgue verte est indiquée par une flèche en pointillés. Les mitochondries et les noyaux des algues ont été éliminés après transfert de gènes nucléaires et plastidiaux (indiqués en bleu, violet et rose) vers le noyau de l'hôte hétérotrophe (N2).



La mitochondrie et le chloroplaste sont les deux seuls organites qui semblent avoir été acquis par endosymbioses. La perte de gènes au cours du temps a réduit les deux symbiontes à des effecteurs métaboliques au sein des cellules hôtes. La cyanobactérie a été progressivement intégrée dans la machinerie cellulaire comme un nouvel organe : le chloroplaste et ainsi domestiquée. L'origine cyanobactérienne des plastes est étayée par des études génomiques et les similitudes ultrastructurales entre les plastes et les cyanobactéries. Le génome cyanobactérien a subi une réduction drastique, la plupart des gènes ont été perdus ou transférés vers le noyau hôte, appelé transfert de gènes endosymbiotique (EGT) (Reyes-Prieto et al., 2007). Le génome des chloroplastes ne code qu'une petite fraction des protéines impliquées dans la photosynthèse (Gould et al, 2008 ; Keeling et al, 2004). Trois lignées auraient évolué à partir de cette première endosymbiose : les Glaucophytes (unicellulaires d'eau douce pourvus de 2 flagelles de longueurs inégales, Chl *a* et phycobiliprotéines), les algues rouges ou Rhodophytes (Chl *a* et phycobiliprotéines) et les algues vertes (Chlorophytes, Chl *a* et Chl *b*). Ces différents groupes sont rassemblés avec les plantes terrestres dans le supergroupe Plantae, aussi nommé Archaeplastida.

L'origine des plastes des autres organismes photosynthétiques est beaucoup plus complexe, elle a fait l'objet ces dernières années de plusieurs études qui s'appuient sur le séquençage de génomes de plus en plus nombreux et des données moléculaires, renforçant ou écartant ainsi certaines hypothèses. Les plastes des algues rouges et des algues vertes auraient ensuite été transférés par endosymbioses secondaires indépendantes (Reyes-Prieto et al., 2007) à des protistes non photosynthétiques qui seraient à l'origine respectivement des Chromalvéolates (Fig. 3B) et des Euglénidés et Chlorarachniophytes amœbæ. L'endosymbiose secondaire à l'origine des Chromalvéolates se serait déroulée à la fin du Précambrien, il y a 0,8 à 1,2 milliards d'années (Simon et al., 2009).

En 1999, Cavalier-Smith a proposé de regrouper dans le supergroupe des Chromalvéolates tous les protistes renfermant de la Chl *c* (Cryptophytes, Haptophytes et Stramenopiles) et les espèces non-photosynthétiques apparantées (Ciliés, Dinoflagellés et Apicomplexa). D'autres analyses phylogénétiques (Hackett et al, 2007 ; Okamoto et Inouye, 2005) ont identifié des membres du groupe des Rhizaria qui ont aussi des affinités avec les Chromalvéolates. Certains Dinoflagellés seraient issus d'une endosymbiose tertiaire au cours de laquelle un Chromiste aurait été phagocyté par un protiste hétérotrophe puis réduit à un chloroplaste (Keeling et al, 2004).

Becker et al. (2008) se sont intéressés à la présence de 39 protéines d'origine *Chlamydiae* chez plusieurs eucaryotes photosynthétiques (Algues et Embryophytes) mais qui sont absentes chez les cyanobactéries. Les *Chlamydiae* sont des bactéries intracellulaires obligatoires des protistes, des invertébrés et des vertébrés mais elles n'ont pas été mises en évidence chez les eucaryotes photosynthétiques. Par conséquent, une endosymbiose de type *Chlamydiae* pourrait avoir infesté le premier eucaryote photosynthétique durant les premiers stades de l'endosymbiose primaire (Fig. 3A), il a plus de 1,1 milliard d'années (Becker et al., 2008). Ces auteurs suggèrent que des gènes ont été acquis par un transfert horizontal entre une bactérie *Chlamydiae* antique et l'ancêtre des eucaryotes photosynthétiques. Par la suite, les gènes codant des protéines *Chlamydiae* se seraient propagés lors des endosymbioses secondaires dans d'autres eucaryotes.

Les transferts de gènes horizontaux ou entre organites et noyau conduisent à des échanges de gènes entre des organismes phylogénétiquement distants, à la création de gènes chimériques et à l'apparition de caractères nouveaux, ces caractères ayant pu disparaître ensuite.

### **1.3. Évolution des diatomées**

Les Straménopiles (ou Hétérokonta) comprennent des parasites tels que l'agent du mildiou (*Phytophthora*) et les diatomées (ou Bacillariophyta) qui sont des algues microscopiques unicellulaires caractérisées par une enveloppe siliceuse externe, ornementée et à structure particulière, le frustule. Des frustules ont été découverts dans des formations géologiques datant du Jurassique supérieur (-146 millions d'années), c'est à dire bien après l'endosymbiose secondaire (Vardi et al., 2008). Les diatomées auraient dominé dans les écosystèmes aquatiques au Tertiaire (-35 millions d'années) et elles pourraient avoir contribué à la formation de gisements de pétrole (Damsté et al., 2004). En se basant sur la symétrie des cellules, ces microalgues ont été classées en deux ordres : les Centriques (ou Médiophycées) à symétrie généralement radiale, connues depuis le Jurassique (-90 millions d'années), et les Pennées (ou Bacillariophycées) à symétrie bilatérale, dont les premiers fossiles retrouvés datent du début du Tertiaire (Round et al., 1990). D'autres critères morphologiques permettent de classer les diatomées en 3 classes : les Coscinodiscophycées (diatomées centriques), les Fragillariophycées (diatomées sans raphé) et les Bacillariophycées (présence d'un raphé) (Round et al., 1990).

Le génome nucléaire des diatomées est une mosaïque de gènes issus d'au moins deux noyaux, celui de la cellule hôte et celui de l'algue rouge, et de la cyanobactérie (Lane et Archibald, 2008). Le séquençage du génome de la diatomée pennée *Phaeodactylum tricornutum* (Bowler et al., 2008) a montré que sa structure était différente de celle du génome de la diatomée centrique *Thalassiosira pseudonana* et qu'environ 40 % des gènes n'étaient pas partagés par les représentants de ces deux lignées qui ont divergé il y a 90 millions d'années. Plus de 500 gènes d'origine procaryotique ont été identifiés dans le génome de *P. tricornutum*, ces gènes proviendraient de protéobactéries, cyanobactéries et d'Archaea. En 2009, Moustafa et al. ont estimé, à partir d'environ 2500 gènes, que plus de 70 % d'entre eux avaient une origine "verte" (Viridiplantae). Comme 85 % de ces gènes sont retrouvés chez tous les Stramenopiles, les auteurs ont avancé l'hypothèse d'un endosymbionte de type Prasinophyte chez l'ancêtre commun des Chromalvéolates, endosymbiose qui se serait donc déroulée avant la phagocytose d'une algue rouge. Tirichine et Bowler (2011) comparent ces endosymbioses à un "ménage à trois" (Fig. 3B). Plus récemment, Deschamps et Moreira (2012) ont contesté l'étude de Tirichine et Bowler (2011) à cause de l'échantillonnage déséquilibré et l'utilisation d'une analyse trop automatisée compte tenu des données disponibles. Après leur nouvelle analyse, Deschamps et Moreira (2012) concluent que les gènes étudiés ont été essentiellement transférés via une algue rouge. Même si certains gènes sont partagés avec la lignée verte, il n'y a pas d'arguments suffisants pour étayer l'hypothèse de l'endosymbiose d'une algue verte.

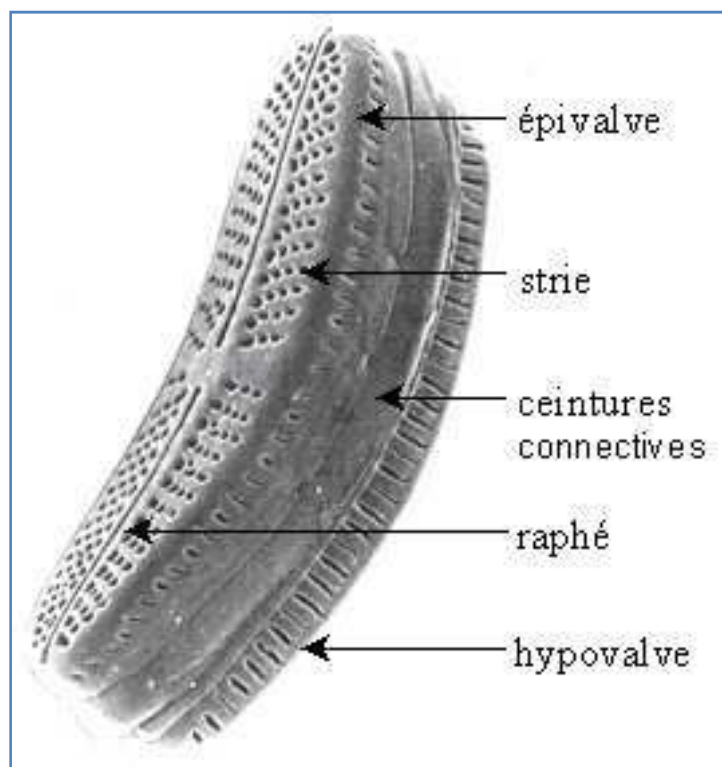
Ces différents travaux montrent que l'évolution des diatomées a suivi une voie fondamentalement différente de celle des plantes, des algues vertes et des autres organismes étroitement apparentés. Toutefois, ces organismes ont un ancêtre commun et ils partagent certaines voies biochimiques similaires, telles que le cycle de l'urée, utilisé à des fins parfois différentes.

## **2. Généralités sur les diatomées**

Les diatomées sont des microalgues unicellulaires, certaines espèces forment des colonies. Présentes dans tous les milieux aquatiques, elles constituent une part majeure du phytoplancton et participent activement à la fixation du  $\text{CO}_2$  et à la production d'oxygène, ce dernier étant indispensable pour la respiration des organismes aérobies.

### **2.1. Morphologie et croissance**

Le frustule siliceux est composé de silice amorphe associée à des protéines et des polysaccharides (Kröger et Sumper, 1998). Il se compose de deux valves (ou thèques) dont les bords sont réunis l'un à l'autre par un nombre variable de bandes intercalaires annulaires (cingulum ou zone connective) (Fig. 4) selon les espèces. La surface des valves et de la zone connective peut être ornée de ponctuations, stries, pores, soies ou protubérances de formes diverses.



**Figure 4 : Frustule d'une diatomée du genre Achnantes (Microscopie électronique à balayage - CRPGL, Luxembourg).**

La multiplication des diatomées se déroule essentiellement par bipartition après une mitose. Le plan de division est parallèle aux valves. Chaque cellule fille conserve une valve de la cellule mère et en sécrète une seconde, plus petite que la première. La taille des cellules diminue donc légèrement lors de chaque division. Dans le milieu naturel, en dessous d'une dimension limite intervient un phénomène compensateur, l'auxosporulation, qui est une reproduction sexuée aux modalités particulières et caractéristiques des Bacillariophyta (Chepurnov et al., 2004 ; Kaczmarska et al., 2013 ; Mouget et al., 2009 ; Round et al., 1990).

## **2.2. Rôle écologique et valorisation**

Les océans couvrent environ 70 % de la surface de la Terre, ils représentent le plus grand écosystème terrestre et leur production nette annuelle représente 40 % de la production globale de la planète (Falkowski et al., 1998). Les diatomées qui représentent environ 40 % du phytoplancton marin, constituent l'un des taxons les plus diversifiés (Round et al., 1990 ; Wilhelm et al., 2006). Bien qu'elles ne constituent qu'une petite fraction de la biomasse photosynthétique de notre planète, elles pourraient fixer, selon Mann (1999), 20 Pg (= 20. 10<sup>9</sup> t) de carbone par an, soit un peu plus que toutes les forêts tropicales (17,8 Pg) qui constituent l'un des écosystèmes les plus productifs terrestres (Field et al., 1998). Les diatomées forment la majeure partie du plancton dans les zones riches en nutriments (généralement les zones côtières océaniques) et les mers froides (Mock et Valentin, 2004; Smetacek, 1999) où elles jouent un rôle prépondérant dans les cycles biogéochimiques (notamment C, Si, O) (Raven et Waite, 2004; Smetacek, 1999) et comme producteurs primaires des chaînes alimentaires.

Les diatomées peuvent être responsables de "blooms" lorsqu'elles prolifèrent (jusqu'à plusieurs millions de cellules/mL) en particulier au printemps. Des espèces du genre *Pseudo-nitzschia* (*P. pseudodelicatissima*, *P. multiseriata*) sont connues pour leur production d'acide domoïque qui est une neurotoxine responsable d'empoisonnements amnésiques aux fruits de mer (Ifremer, 2006). Ces espèces peuvent avoir un effet nuisible sur l'écosystème local, la pêche, l'aquaculture et le tourisme. Par ailleurs, les diatomées contribuent largement au "biofouling" qui correspond à l'accumulation de microorganismes, algues et animaux sur des surfaces immergées, du fait de leur capacité à adhérer à des surfaces et à produire des exopolysaccharides mucilagineux.

Le biofouling pose d'énormes problèmes de maintenance des coques des bateaux, des filets et cages de pêche, et des installations de désalinisation. Les traitements antifoulings sont à base de substances biocides qui s'accumulent dans le milieu marin et intoxiquent un large spectre d'organismes.

Une meilleure connaissance de l'écologie des diatomées a permis leur utilisation généralisée en tant qu'indicateurs biologiques de la qualité des eaux et de l'état des écosystèmes. La rapidité de leur cycle de développement et leur sensibilité aux pollutions, notamment organiques, azotées et phosphorées en font des organismes intéressants pour la caractérisation de la qualité d'un milieu. Un indice biologique diatomique (IBD) variant de 1 (eau polluée) à 20 (eau pure) est ainsi utilisé pour évaluer la qualité des cours d'eau.

La sédimentation constante des frustules de diatomées mortes a conduit à la formation de couches sédimentaires particulièrement importantes dans des régions océaniques. Il y a ainsi des dépôts fossiles datant du Tertiaire appelés "diatomites" qui sont exploitées pour fabriquer des filtres, des déodorants, des agents décolorants, des agents abrasifs pour dentifrice, etc.

- Les diatomées sont cultivées à grande échelle pour nourrir des animaux marins d'élevage (aquaculture) tels que des mollusques (huîtres, moules, etc), poissons, crustacés ou des larves (Müller-Feuga, 1997).

- Les diatomées sont aussi de plus en plus produites et commercialisées comme compléments alimentaires (exemple : *Odontella aurita*). En effet, elles sont notamment une source d'acides gras polyinsaturés à longue chaîne de la série oméga-3 tels que l'EPA (acide eicosapentaénoïque, 20:5 $\omega$ 3) et le DHA (acide docosahéxaénoïque, 22:6 $\omega$ 3) (Domergue et al., 2002; Guihéneuf et al., 2008). Ces acides gras sont essentiels et possèdent des propriétés bénéfiques sur la santé puisqu'ils favorisent le développement du système nerveux et préviennent les risques cardio-vasculaires (Dyerberg et Jorgensen, 1982 ; Lagarde et al., 1986 ; Mimouni et al., 2012 ; Rousseau et al., 2003). Les poissons sont la source habituelle d'oméga 3 mais cette ressource est en cours d'épuisement et les polluants s'accumulent dans les graisses des animaux. Les microalgues constituent ainsi une source alternative aux huiles de poissons, renouvelable et elles peuvent être cultivées en systèmes clos. Elles synthétisent des

pigments caroténoïdes dont la fucoxanthine qui présente un fort potentiel thérapeutique du fait de ses activités cytotoxique, antiproliférative et proapoptotique (Ishikawa et al., 2008). D'autres molécules synthétisées par les diatomées comme des antioxydants pourraient être valorisées.

- Certaines microalgues peuvent produire de grandes quantités de lipides (jusqu'à 70 % de la masse sèche chez *Botryococcus braunii* et *Nannochloropsis* sp.) qui s'accumulent sous forme de gouttelettes dans le cytoplasme, notamment en condition de stress (Chisti, 2007). Certaines diatomées comme *Phaeodactylum tricornutum* ou *Nitzschia* sp. renferment jusqu'à 30 et 47 % respectivement de leur masse sèche en lipides. Chisti (2007) estime que les microalgues pourraient être une source renouvelable et modulable de biodiésel qui ne serait pas en concurrence avec des oléagineux alimentaires.

- Dans le secteur des nanotechnologies, les motifs microscopiques des frustules servent de modèles pour des semi-conducteurs et des nanopuces. L'industrie pharmaceutique développe aussi des nanocapsules inspirées de la structure des frustules pour diffuser plus lentement des principes actifs dans le corps des malades.

La transformation génétique des diatomées est utilisée à des fins de recherche mais elle pourrait aussi permettre d'augmenter leur valorisation. En effet, des microalgues servent déjà à produire des protéines recombinantes (Potvin et Zhang, 2010).

### **2.3. Photosynthèse oxygénique**

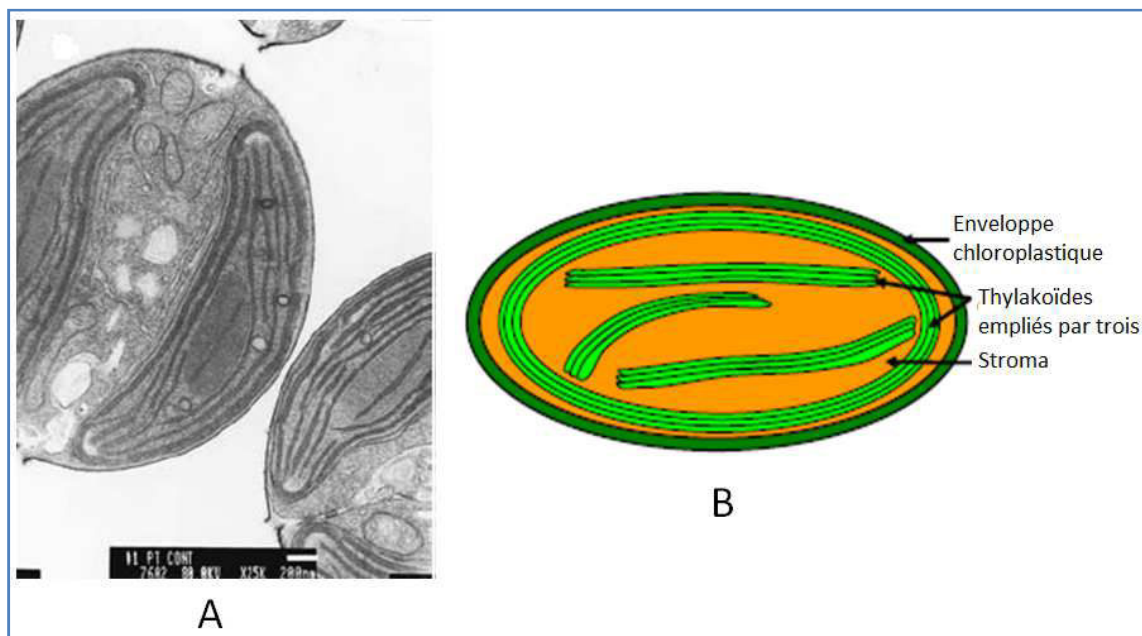
Les diatomées, comme les autres algues et les plantes, élaborent en présence d'énergie lumineuse leur propre matière organique à partir de l'eau, du CO<sub>2</sub> et des sels minéraux. L'ensemble des réactions impliquées dans la synthèse des premières molécules organiques glucidiques constitue la photosynthèse. Cette dernière se déroule dans les chloroplastes où la chlorophylle capte l'énergie lumineuse et la transforme en énergie chimique.

#### **2.3.1. Chloroplastes**

Les diatomées renferment des chloroplastes dont, selon les espèces, le nombre varie de un à plusieurs dizaines et qui peuvent avoir une forme de grains, de plaques, de bandes, voire d'étoiles. Les plastes des diatomées renferment de la Chl *a* et de la Chl *c* (*c1* et *c2*), ainsi que des pigments caroténoïdes ( $\beta$ -carotène, fucoxanthine, diadinoxanthine) qui

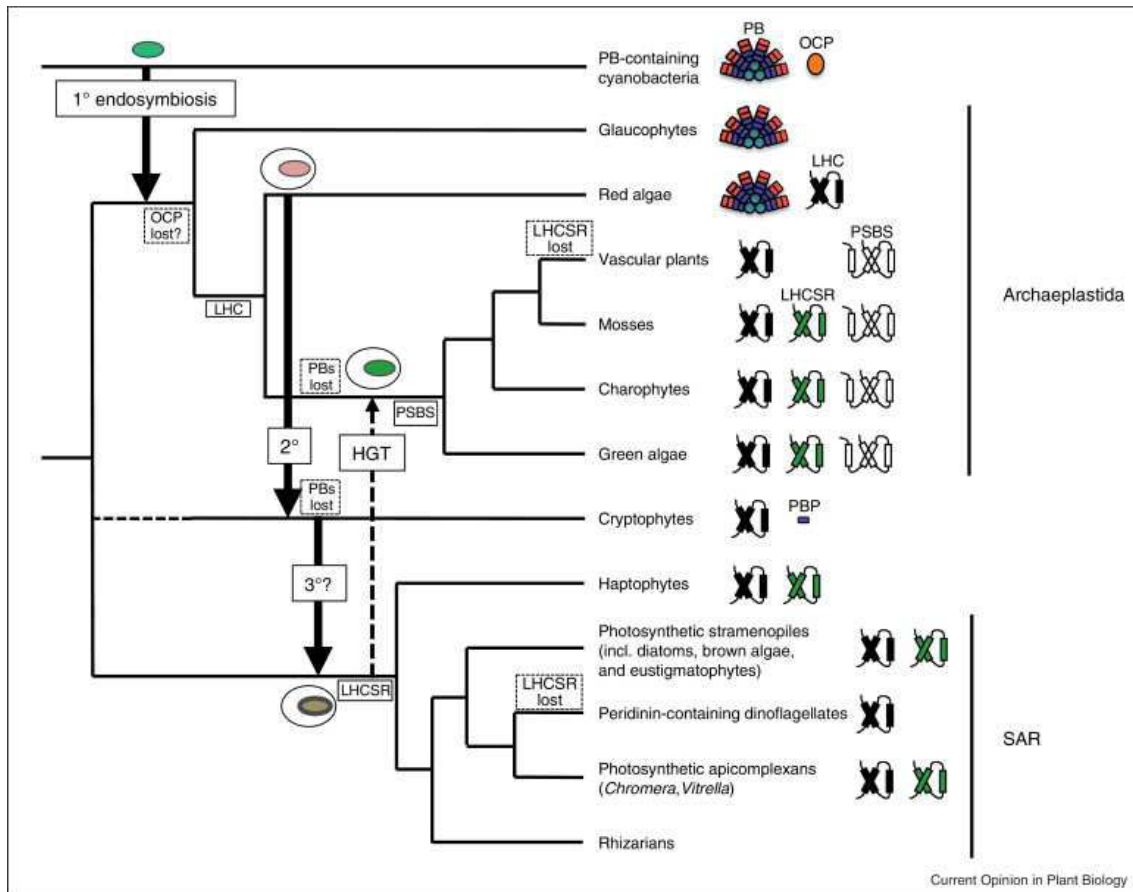
masquent plus ou moins la couleur verte des chlorophylles. A la différence de ceux des plantes terrestres, les chloroplastes des diatomées ne sont pas verts mais jaunes ou bruns.

Les différentes endosymbioses ont conduit à la formation de chloroplastes délimités par un nombre de membranes variant de deux chez les Archeplastidia et algues rouges à quatre chez les diatomées. Chez ces dernières, les deux membranes les plus internes de l'enveloppe plastidiale correspondent à l'enveloppe plastidiale du "chloroplaste de l'algue rouge", la troisième membrane dériverait de la membrane plasmique de l'algue rouge et la membrane la plus externe, recouverte de ribosomes est en continuité avec le réticulum endoplasmique de la cellule hôte (Armbrust et al., 2004, Wilhelm et al., 2006). Les thylacoïdes (sacs membraneux) sont groupés par trois et trois thylacoïdes accolés, formant la "guirlande" typique de ce groupe, doublent l'enveloppe plastidiale (Fig. 5B) (Jupin et Lamant, 1997). Il est à remarquer que tous les Chromalveolates photosynthétiques renferment de la Chl *c* qui est absente chez les algues rouges. Ce caractère très original résulterait de la création de gènes chimériques suite à des transferts de gènes horizontaux ou des transferts entre organites et le noyau de la cellule hôte.



**Figure 5 : A-Coupe dans des cellules de *Phaeodactylum tricornutum* renfermant deux chloroplastes observés en microscopie électronique (d'après Falkowski et Raven, 1997). B-Schéma de la structure d'un chloroplaste de diatomée avec des thylakoïdes empilés par trois.**





**Figure 6 : Schéma des relations entre les principaux groupes d'organismes photosynthétiques oxygénés, les étapes évolutives possibles et leurs types d'antennes collectrices de la lumière. D'après Niyogi et Truong(2013).**

La longueur des branches sont dessinées pour des commodités de lecture et ne sont pas destinées à indiquer des longueurs spécifiques de temps. Les endosymbioses hypothétiques, les transferts horizontaux de gènes (HGT) et les événements d'acquisition sont présentés dans des cadres en trait plein; les pertes sont présentées dans des cadres en pointillés. Les Cryptophytes ont une antenne inhabituelle de phycobiliprotéines (PBP) qui ne font pas partie d'un phycobilisome (PB). OCP : complexe protéine-caroténoïde orange; LHC : *Light-harvesting* complexe ; LHCSR : protéine spécifique de LHC induite par le stress; PSBS : protéine du complexe LHC ; SAR : Stramenopiles, Alveolates et Rhizaria.

A l'intérieur des chloroplastes, des molécules d'ADN, des ribosomes (ARN) et des globules lipidiques sont aussi présents. Un ou plusieurs pyrénoides (Nassiri et al., 1998; Tachibana et al., 2011) peuvent être fixé(s) sur les thylacoïdes, ils renferment de la RuBisCO (50 à 99 % de la RuBisCO totale selon la teneur en CO<sub>2</sub> dans le milieu ambiant, Moroney et Somanchi, 1999) et de l'anhydrase carbonique (Tachibana et al., 2011). Les diatomées ne produisent pas d'amidon, le stockage de glucides s'effectue dans des vacuoles sous forme de chrysolaminarine qui est un assemblage de β-1-6 glucose et de β-1-3 glucose (Granum et al., 2002).

Au cours de l'évolution des différents organismes photosynthétiques, des changements majeurs sont apparus dans leur système de collecte de la lumière. En effet, les antennes collectrices de la lumière constituent le paramètre le plus variable de l'appareil photosynthétique (Fig.6), indiquant l'importance adaptative et écologique de l'absorption de la lumière dans les différents environnements (Niyogi et Truong, 2013).

### **2.3.2. Photosynthèse**

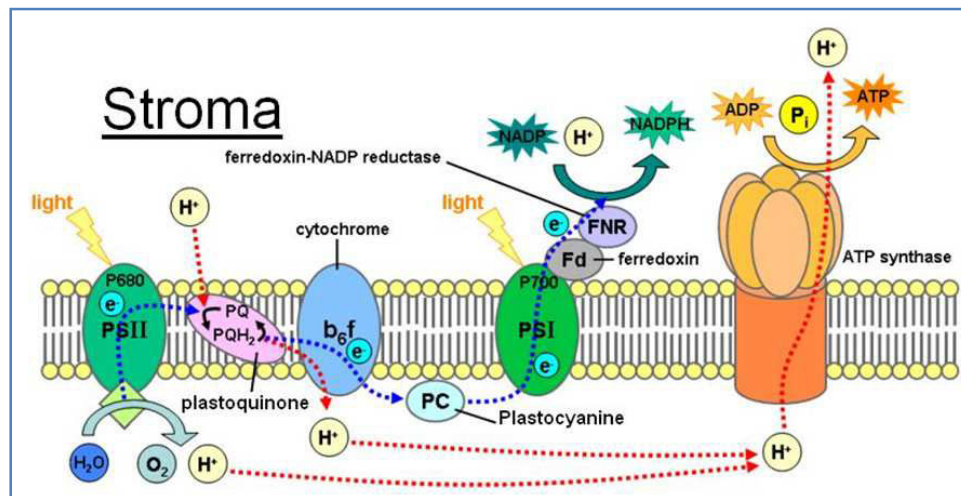
La photosynthèse fait intervenir deux séries de réactions : les réactions photochimiques correspondant à la phase claire durant laquelle la scission de molécules d'eau libère du dioxygène (O<sub>2</sub>) et les réactions sombres correspondant au cycle de Calvin-Benson durant lequel les premiers oses sont synthétisés.

Réaction globale de la photosynthèse :



La photosynthèse oxygénique nécessite la coopération de deux photosystèmes (PS) nommés PSI et PSII (Fig. 7) composés de pigments photorécepteurs et de protéines. Chez les diatomées, chaque PS est associé à une antenne collectrice de la lumière composée de Chl *a* et *c*, de pigments caroténoïdes et de protéines transmembranaires (Wilhelm et al., 2006). Les pigments caroténoïdes sont essentiellement de la fucoxanthine et de la diadinoxanthine d'où leur nom "Light harvesting Complex with Fucoxanthin" (LHCF) qui permettent une absorption dans la région bleu-vert du spectre visible. Chez les algues vertes et les cyanobactéries, les PSI sont aussi nombreux que les PSII. Chez les diatomées, le nombre de PSI est fréquemment plus faible que celui des PSII probablement à cause d'une carence en fer, le PSI renferme quatre fois plus d'atomes de fer que le PSII (Strzepeck et Harrison,

2004) mais il pourrait aussi s'agir d'une adaptation à de forts éclaircements (Jupin & Lamant, 1997).



**Figure 7 : Schéma de l'organisation de la membrane d'un thylakoïde.**

Lorsque les diatomées sont éclairées, la lumière solaire est utilisée comme source d'énergie pour réaliser la photolyse de l'eau, ce qui se traduit par un dégagement d' $O_2$ . L'absorption de photons provoque l'excitation transitoire de molécules de Chl *a* situées dans le centre réactionnel du PSII. Le retour à l'état fondamental peut s'effectuer selon trois voies complémentaires (Fig. 8) : libération d'un électron (séparation de charge conduisant à la photochimie), émission d'un photon de grande longueur d'onde (fluorescence) ou dissipation thermique. Le flux d'électrons entre les différents transporteurs membranaires s'accompagne de la production d'ATP et de NADPH (Fig. 7) qui serviront à la fixation de  $CO_2$ . La principale voie de fixation du carbone inorganique chez les diatomées est comparable à celle existant chez les plantes C3. Elle implique la ribulose biphosphate carboxylase oxygénase (RuBisCO) qui catalyse la fixation du  $CO_2$  sur une molécule de ribulose biphosphate, l'hexose formé se scinde en deux acides phosphoglycériques qui sont à l'origine de toutes les molécules synthétisées. Plusieurs auteurs ont montré que la fixation du carbone pouvait aussi être assurée, chez quelques espèces, par une voie de type C4-like catalysée par la phosphoénolpyruvate carboxylase (PEPcase) dans le cytosol (Rech et al, 2008 ; Reinfelder et al, 2000 ; Roberts et al, 2007).

Les diatomées sont responsables d'au moins un quart du carbone inorganique (Ci) fixé chaque année dans l'océan (Granum et al., 2005) bien que la teneur en  $CO_2$  dans l'eau de mer soit très faible, 10 à 25  $\mu M$  ; le Ci étant essentiellement sous forme d'ions

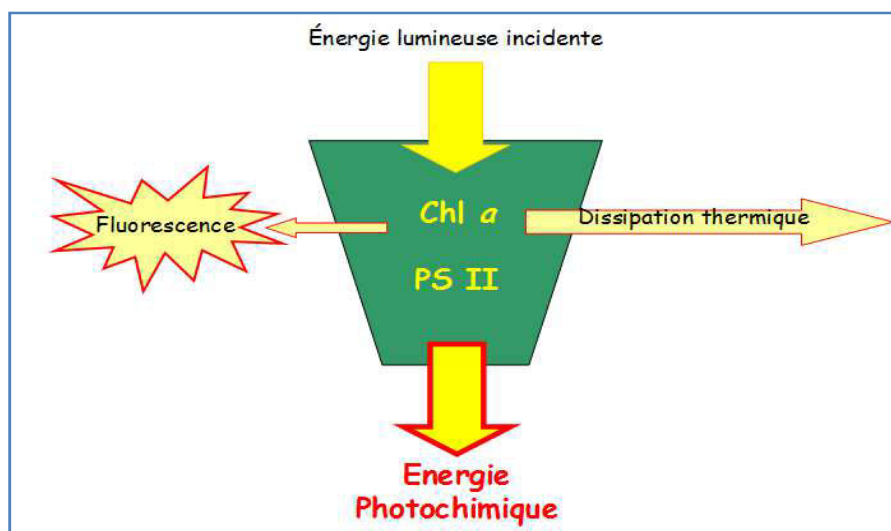
bicarbonates. Avec une constante de Michaelis pour la RuBisCO comprise entre 30 et 40  $\mu\text{M}$  (Badger et al., 1998), l'affinité de cette enzyme vis à vis du  $\text{CO}_2$  chez des diatomées est parmi les plus élevées de tout le phytoplancton (Tortell, 2000). Dans ce milieu limitant en substrat ( $\text{CO}_2$ ) pour la RuBisCO, les diatomées, comme la plupart des microalgues, ont développé des mécanismes de concentration du  $\text{CO}_2$  à proximité de la RuBisCO (Wilhelm et al., 2006). Parmi ceux-ci, l'anhydrase carbonique qui catalyse la déshydratation réversible des ions bicarbonates en  $\text{CO}_2$  joue un rôle majeur. Cette métallo-enzyme renferme un atome de zinc au niveau de son site catalytique mais cet élément peut être remplacé par un atome de cobalt ou de cadmium (Price et Morel, 1990). Park et al. (2007) suggèrent que cette faculté de remplacement d'un élément par un autre a permis aux diatomées de devenir les producteurs primaires majoritaires dans les océans où les concentrations en  $\text{CO}_2$  sont faibles. Les autres mécanismes de concentration du  $\text{CO}_2$  consistent en une absorption active de celui-ci et des ions  $\text{HCO}_3^-$  et l'accumulation de  $\text{CO}_2$  grâce à une voie de type C4-like (voir la synthèse de Giordano et al., 2005).

Les diatomées ont besoin de lumière pour photosynthétiser, mais elles sont soumises à un éclaircissement fluctuant en quantité et en qualité lorsqu'elles sont dans la colonne d'eau. Généralement, les diatomées ont une meilleure croissance sous faible éclaircissement (Smetacek, 1999). Le suivi de la croissance de quelques espèces à différents niveaux d'éclaircissement au laboratoire montre que l'éclaircissement est saturant entre 40 et 375  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  selon les espèces et les conditions de culture (Morant-Manceau et al., 2007 ; Mouget et al., 1999; Nguyen-Deroche et al., 2012 ; Watson et al., 2004).

### **2.3.3.Fluorescence de la chlorophylle *a***

Chez les plantes saines, environ 80 % de l'énergie lumineuse collectée par les molécules de Chl *a* vont servir à la synthèse de glucides, la fluorescence représente entre 2 et 5 % et le reste est dissipé sous forme de chaleur. A température ambiante, la fluorescence émise par la Chl est due essentiellement à la Chl *a* située dans le centre réactionnel du PSII (Govindjee, 1995). Les trois voies de dé-excitation de la Chl *a* co-existent de manière compétitive (Fig. 8). Ainsi, la mesure des différents paramètres de la fluorescence de la Chl *a* permet de connaître l'état physiologique du PSII, l'efficacité des réactions photochimiques et d'estimer la perte d'énergie sous forme de chaleur (Maxwell et Johnson, 2000). Les mesures de fluorescence présentent l'avantage d'être

rapides et non intrusives. Elles sont largement utilisées pour suivre les mécanismes d'acclimatation et d'adaptation des organismes photosynthétiques soumis à différents stress environnementaux tels que de forts éclaircements, les UV, la sécheresse, des variations de température, etc (Nguyen-Deroche et al., 2012 ; Rech et al., 2005 ; Rohacek et al., 2008).



**Figure 8 : Schéma des voies de dissipation de l'énergie lumineuse au niveau du photosystème II.**

Lorsque l'énergie ne peut pas être transmise par la voie photochimique (transmission d'un électron au Photosystème I), elle peut être dissipée selon deux modes : la chaleur ou la fluorescence. Chl *a* : Chlorophylle *a*, PSII : Photosystème II.

L'efficacité photosynthétique et par conséquent la production de biomasse dépend fortement du niveau d'éclairement qui est très variable en quantité et en qualité dans la colonne d'eau. Les forts éclaircements en particulier peuvent endommager les PSII (Külheim et al., 2002). Les diatomées possèdent des mécanismes de protection de leur appareil photosynthétique parmi lesquels le quenching non-photochimique (NPQ) (ou atténuation non-photochimique de la fluorescence) occupe un rôle majeur. Deux processus dans le NPQ permettent de dissiper rapidement l'excès d'énergie, il s'agit de la dissipation d'énergie sous forme thermique et le cycle des xanthophylles autour du PSII et/ou PSI (Lavaud et al ; 2002a ; 2002b ; 2002c). Ce cycle implique la diadinoxanthine qui possède la capacité de dissiper l'énergie lumineuse en excès grâce à une série de réactions de dé-époxydation et époxydations cycliques. La diadinoxanthine est dé-époxydée en diatoxanthine grâce à une diadinoxanthine dé-époxydase (DDE) sous fort éclairciment, libérant ainsi de l'eau. La diatoxanthine est ensuite re-oxydée en diadinoxanthine à l'obscurité ou sous faible éclairciment.

Récemment, les travaux de Bailleul et al. (2010) ont mis en évidence le rôle clé d'une protéine de stress de l'antenne collectrice de lumière, nommée LHCX1 (Light-Harvesting complex stress-related protein) chez *P. tricornutum*, qui amplifie le NPQ. La présence constitutive de cette protéine dans les cellules non stressées permettrait aux diatomées de réagir très rapidement aux variations d'éclairement. L'efficacité de ces mécanismes est variable selon les espèces et expliquerait le succès écophysiologique de certaines microalgues (Lavaud et Lepetit, 2013 ; Wilhelm et al., 2006).

#### **2.4. Génétique et génome des diatomées**

Le génome complet de la diatomée centrique *Thalassiosira pseudonana* et celui de la diatomée pennée *Phaeodactylum tricornutum* ont été séquencés et sont disponibles dans la banque de données Joint Genome Institute. Ces génomes sont de petite taille, 32,4 Mb pour *T. pseudonana* et 27,4 Mb pour *P. tricornutum* (Armbrust et al., 2004 ; Bowler et al., 2008), comparés à celui de plantes (115 Mpb pour *Arabidopsis thaliana* et 68000 Mpb pour le pin) ou de microalgues vertes (112 Mpb pour *Chlamydomonas reinhardtii*) mais de plus grande taille que celui de la microalgue rouge *Cyanidioschyzon merolae* (16 Mpb). L'analyse de ces deux génomes et de très nombreuses séquences exprimées (Expressed Sequence Tag) a donné lieu à plusieurs publications. *P. tricornutum* renferme 10402 gènes et *T. pseudonana* 11776, 57 % des gènes de *P. tricornutum* sont partagés avec *T. pseudonana* et 1328 sont absents des autres génomes eucaryotes séquencés (Bowler et al., 2008). La découverte chez *T. pseudonana* des gènes d'un cycle complet de l'urée était inattendue (Armbrust et al., 2004) ; de nouveaux gènes impliqués dans le transport de l'acide silicique et la formation de la paroi silicifiée, l'absorption du fer, la synthèse d'acides gras polyinsaturés et l'utilisation de composés azotés ont été identifiés.

Les alignements de séquences ont aussi permis d'identifier 171 gènes d'origine "algue rouge" chez *P. tricornutum* dont 108 sont aussi présents chez *T. pseudonana*. Parmi ces gènes, 11 sont présents chez les oomycètes (*Phytophthora*) qui seraient issus, comme les diatomées, d'une endosymbiose secondaire d'une algue rouge mais qui auraient ensuite perdu le chloroplaste. Ce résultat conforte l'hypothèse de l'endosymbiose secondaire, du transfert de gènes entre le noyau de l'algue rouge et celui de la cellule hôte et de la proximité phylogénétique des diatomées avec les oomycètes (Armbrust et al., 2004 ; Bowler et al., 2008). Par ailleurs, le génome des deux diatomées partage de nombreux gènes (plus de 580 chez *P. tricornutum*) avec différents groupes

deprocaryotes conférant ainsi de nouvelles capacités métaboliques. Des transferts de gènes horizontaux entre procaryotes et eucaryotes se seraient produits après l'endosymbiose secondaire et auraient apporté 5 % du génome de *P. tricornutum* (Bowler et al., 2008).

Une faible partie de ces génomes est constituée d'éléments génétiques mobiles (ou transposables) : 1,1 % chez *T. pseudonana* et 5,8 % chez *P. tricornutum*. Ces éléments transposables (ETs) sont essentiellement des rétrotransposons appartenant à la famille *Ty1-copia* qui représente 90 % des ET chez *P. tricornutum* (Bowler et al., 2008 ; Maumus et al., 2009). La partie "éléments transposables" sera développée une autre partie dans la page 36.

Plus récemment, le génome de la diatomée pennée psychrophile *Fragilariopsis cylindrus* (81 Mpb) a été séquencé (<http://genome.jgi-psf.org/Fracy1/Fracy1.home.html>). Son analyse devrait permettre de mieux comprendre l'adaptation de cette espèce au milieu polaire. Actuellement, le génome de la diatomée *Pseudonitzschia multiseries* produisant des peptides neurotoxiques, est en cours de séquençage (<http://genome.jgi.doe.gov/Psemu1/Psemu1.home.html>). Cette espèce provoque à certaines périodes de l'année des blooms dans les eaux côtières, pouvant provoquer des intoxications.

Ces quatre génomes de diatomées fournissent une base pour des analyses comparatives et de nouveaux éléments moléculaires nécessaires à la compréhension du succès écologique des diatomées dans les océans. Ils pourraient aussi faciliter l'utilisation de ces organismes comme "usines" potentielles pour l'expression de gènes d'intérêt pour l'industrie (Potvin et Zhang, 2010 ; Mimouni et al., 2012 ; Stengel et al., 2011).

## **2.5. Réponses des diatomées aux contraintes de l'environnement**

### **2.5.1. Définition**

Tout changement d'un ou plusieurs facteurs environnementaux a un impact sur la croissance et le développement des organismes végétaux, mais les déviations d'un paramètre par rapport à son optimum ne sont pas nécessairement un stress pour un organisme tolérant et adapté à son milieu (Gaspar et al., 2002). Hoffmann et Parsons (1997) ont défini un stress comme "tout changement environnemental qui réduit considérablement la bonne santé de l'organisme". Les conditions défavorables dues à des facteurs biotiques ou abiotiques affectent directement ou indirectement la

physiologie des plantes en altérant leur métabolisme, leur croissance et leur développement. Le stress est d'abord perçu par les récepteurs puis le signal est transformé en une cascade de signaux biochimiques qui pourront modifier l'expression de gènes et induire l'expression de gènes de défense pour répondre à court terme par des changements morphologiques, biochimiques et physiologiques. Les mécanismes qui permettent de survivre aux stress sont appelés mécanismes de résistance. Lorsque les dommages sont plus importants que les processus de réparation, ils peuvent provoquer la mort cellulaire.

Bijlsma & Loeschcke (1997) ont identifié deux types de réponse aux stress :

- l'acclimatation est une réponse physiologique individuelle pour s'accommoder aux changements de l'environnement (contexte individuel)
- l'adaptation résulte de modification génétique et de sélection de génotypes favorables dans une population (contexte évolutif).

### **2.5.2. Réponses des diatomées marines aux stress**

Dans les conditions naturelles, les diatomées marines sont particulièrement abondantes dans les zones côtières riches en nutriments. En haute mer, la teneur en fer et autres microéléments est faible, ce qui réduit fortement la densité des microalgues eucaryotes. Toutefois, si le silicium n'est pas limitant, les diatomées peuvent se développer à des concentrations en fer de quelques nmoles L<sup>-1</sup> car elles présentent une forte affinité vis à vis de cet élément et qu'elles sont capables de réduire leur besoin en fer (Falkowski et al., 2004 ; Sunda et Huntsman, 1997 ; Wilhelm et al., 2006). Des teneurs élevées en microéléments tels que le Cu (3,15 µM) ou le Zn (20,25 µM) affectent la croissance et l'activité photosynthétique de diatomées marines telles que *A. acutiuscula* (Cid et al., 1995 ; Nguyen-Deroche et al., 2009 ; Nguyen-Deroche et al., 2012 ; Stauber & Florence, 1990). Outre les éventuelles carences (Fe, CO<sub>2</sub>, lumière, etc) qui peuvent nuire au bon développement des diatomées, elles peuvent être, par ailleurs, exposées à des facteurs externes en excès devenant toxiques (nutrients, lumière, etc) et à des facteurs nocifs (UV, pesticides, etc). Ainsi, l'étude de la microflore de marais salants a montré que l'abondance et la diversité des diatomées diminuent le long du gradient de salinité ; par exemple, les espèces les plus résistantes sont encore présentes à une salinité de 190 dans la saline de Sfax en Tunisie (Abid et al., 2008).

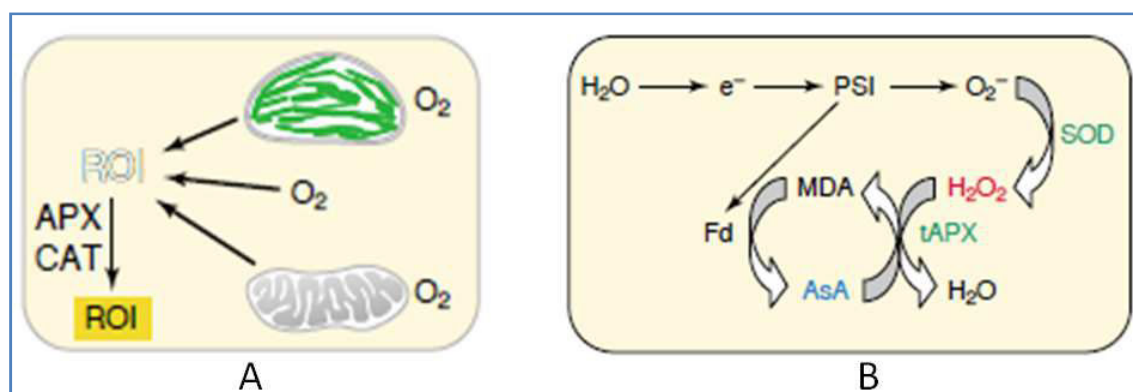
Les diatomées sont présentes à toutes les latitudes, leur répartition qui s'étend des pôles (exemple : *Fragilariopsis cylindrus*) aux régions tropicales (*A. acutiuscula*) en passant



par les zones tempérées (*Skeletonema costatum*) indique leur capacité d'adaptation (Mock et Valentin, 2004). Cultivé au laboratoire, *P. tricornutum* présente une croissance optimale à 21-22 °C ainsi qu'une production d'acides gras maximale (Yongmanitchai et Ward, 1991). Toutefois, les acides gras présentant un intérêt nutritionnel sont polyinsaturés. La teneur en acides gras polyinsaturés (AGPI) chez *P. tricornutum* diminue lorsque la température augmente entre 10 et 25 °C (Jiang et Gao, 2004). De même, lorsque les cellules cultivées à 25 °C sont transférées à une température plus faible (20, 15 ou 10 °C), la synthèse d'AGPI dont l'acide écosapentaénoïque (EPA), est augmentée. Par ailleurs, Chen (2012) a montré que chez 12 espèces de diatomées marines cultivées sous serre, la composition en AGPI et en acides gras saturés variait différemment selon l'espèce, le niveau d'éclairement et la température (17 à 32 °C). Les études réalisées sur les plantes terrestres montrent que les basses températures rigidifient les membranes (fluidité diminuée), modifient leur composition en augmentant le degré d'insaturation des acides gras composant la bicouche phospholipidique (Uemura et al., 2006) et réduisent l'activité de nombreuses enzymes. Les variations de température ont ainsi un impact sur la photosynthèse au niveau des réactions qui ont lieu dans les membranes des thylacoïdes et au niveau des activités enzymatiques (Rai et Gaur, 2001 ; Saakov, 2002; Rohacek et al., 2008). Chez les plantes, l'expression des gènes codant des facteurs de transcription de type CBF (C-repeat Binding Factor) est rapidement induite (15 min) par une température basse et elle atteint un maximum après 2 h d'induction (Stockinger et al., 1997 ; Van Buskirk et Thomashow, 2006). Les protéines CBF produites se fixent ensuite sur des promoteurs et induisent ou répriment la transcription de gènes cibles impliqués dans la réponse au froid (Van Buskirk et Thomashow, 2006). Chez la diatomée *Fragilariopsis cylindrus*, la protection contre la formation de cristaux de glace dans les cellules est assurée par des protéines "anti-gel" comme dans d'autres organismes adaptés au climat polaire (Bayer-Giraldi et al., 2011). Les gènes codant ces protéines auraient été acquis par transfert horizontal entre des basidiomycètes et le genre *Fragilariopsis* (Sorhannus, 2011).

L'étude des protéines de plantes soumises à des contraintes environnementales, notamment des températures élevées, a révélé la synthèse accrue de protéines chaperonnes nommées HSP (Heat Shock Proteins) dont le rôle est de prévenir la dénaturation des autres protéines (Vierling, 1991). Elles interviennent également lors d'autres stress tels que la sécheresse ou le froid qui induisent un stress oxydant (Banzet et al., 1998; Van Berkel et al., 1994). Ces protéines sont classées en cinq groupes selon

leurs masses moléculaires : les HSP 100, 90, 70, 6 et les HSP de faible masse moléculaire sHSP (small HSP) (Vierling, 1991). Armbrust et al. (2004) ont trouvé dans le génome de *T. pseudonana* des séquences codant des facteurs de transcription essentiellement de la famille des HSP, ce qui est relativement rare chez les autres eucaryotes. Les sHSP ont été recherchées dans cinq génomes séquencés de microalgues *Chlamydomonas reinhardtii*, *Cyanidioschyzon merolae*, *Ostreococcus lucimarinus*, *Ostreococcus tauri*, and *Thalassiosira pseudonana* (Waters et Rioflorido, 2007), il ressort de cette étude que le nombre et la diversité des sHSP ne sont pas corrélés avec l'adaptation à des conditions extrêmes.



**Figure 9 : A- Sites de production des formes réactives de l'oxygène (ROI) dans une cellule végétale ; B- Production d'ion superoxyde ( $O_2^-$ ) dans le chloroplaste d'une plante et voie de détoxification enzymatique. SOD : superoxyde dismutase ; APX : ascorbate peroxydases ; CAT : catalase. D'après Mittler (2002).**

Les conditions défavorables réduisent la croissance des diatomées en perturbant leur métabolisme et en générant, dans la grande majorité des cas, un stress oxydatif (Bertrand et Poirier, 2005 ; Nguyen-Deroche et al, 2012; Mittler, 2002 ;Pinto et al., 2003 ; Rijstentbil, 2003 ; Rijstenbil et al., 1994). Celui-ci est dû à une augmentation de la production de formes réactives de l'oxygène ( $O_2^-$ ,  $OH^\cdot$ ,  $H_2O_2$ , etc) dans les différents compartiments cellulaires qui perturbent le métabolisme en oxydant certaines molécules comme les lipides par exemple. Les formes réactives de l'oxygène (Reactive Oxigen Species) ou radicaux libres sont des produits dérivés de la respiration et de la photosynthèse (Fig. 9A). Les organismes sont pourvus de mécanismes de défense contre les ROS ; il s'agit d'antioxydants tels que l'ascorbate, l' $\alpha$ -tocophérol et d'antioxydants enzymatiques tels que la superoxyde dismutase (SOD) ou la catalase (Fig. 9). Quatre SOD différant par leur cofacteur métallique ont été identifiées: Cu/Zn-SOD, Fe-SOD, Ni-SOD et Mn-SOD, cette dernière est présente chez les diatomées (Wolfe-Simon et al.,

2005). Les SOD sont localisées dans les chloroplastes et les mitochondries, ce sont les premières enzymes à intervenir dans l'élimination des anions superoxydes en catalysant leur dismutation en peroxyde d'hydrogène qui est ensuite réduit en eau par l'ascorbate peroxydase.

### **3.Éléments transposables**

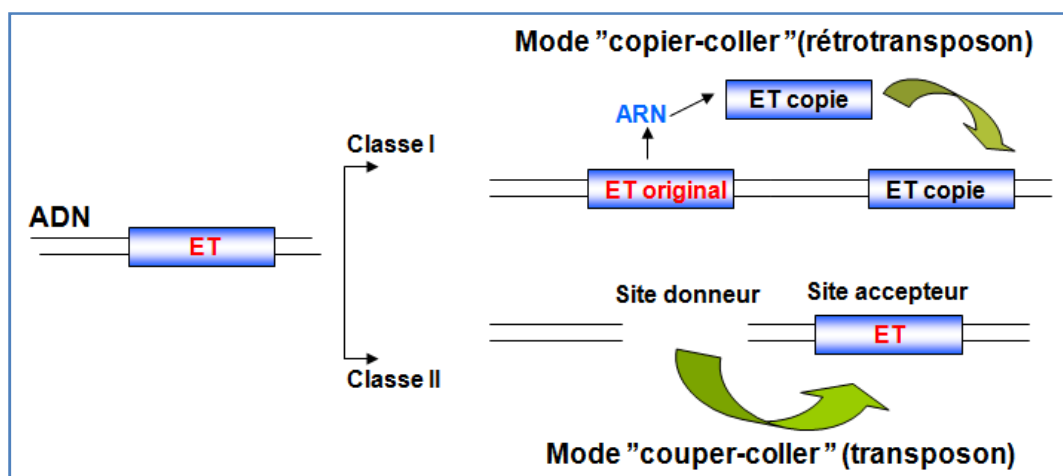
Les génomes ont longtemps été considérés comme des entités stables avec des positions quasiment immuables des séquences et donc des gènes. Dès les années 1950, McClintock mettait en évidence la présence des séquences mobiles chez le maïs qui sont responsables de la création de mutations qui se manifestent au niveau de la couleur des grains et appelées éléments d'ADN mobile ou éléments transposables (ET). Finalement, ces éléments ont été découverts chez les procaryotes et les eucaryotes. Chaque espèce peut posséder plusieurs types d'éléments transposables et en nombre variable.

Les ET ont d'abord été considérés comme des "parasites génomiques" ou "ADN égoïste" ayant pour seul objectif d'assurer leur survie au sein des génomes. Les ET sont des séquences d'ADN ayant la capacité de se déplacer au sein d'un chromosome ou d'un chromosome à un autre, dans le génome qui les héberge. Ils possèdent dans leur séquence le ou les gènes codant les enzymes nécessaires à leur déplacement encore appelé transposition. On distingue les ET autonomes qui sont mobiles car ils possèdent des gènes codant des protéines fonctionnelles de transposition et les ET non-autonomes qui contiennent des gènes de transposition mutés ou délétés. Toutefois, ces derniers peuvent être mobilisés par des éléments autonomes, ce processus est appelé mobilisation en *trans*.

#### **3.1. Classification**

Les ET ont été initialement classés selon la nature de l'intermédiaire lors de leur transposition, deux mécanismes principaux de déplacement ont été décrits (Finnegan, 1989). Le premier niveau de classification est composé de deux classes, les éléments de classe I et de classe II (Finnegan, 1989). La classe I regroupe tous les rétrotransposons qui sont des éléments qui transposent selon un mécanisme de type "copier-coller" (Fig. 10). La transposition s'effectue au moyen de la production d'un ARN rétrotranscrit en ADN complémentaire (ADNc). Celui-ci correspond à une copie du rétrotransposon qui est ultérieurement intégrée dans un locus accepteur du génome. Les ET de classe II, ou

transposons, transposent directement, sans intermédiaire, selon un mécanisme "couper-coller" (Fig. 10). Ainsi l'élément présent au sein du site donneur est excisé et réinséré dans le génome dans un site accepteur.



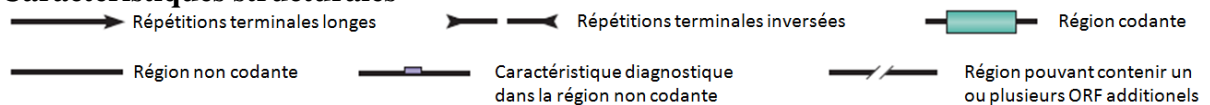
**Figure 10 : Schéma de la classification des ET selon leur mode de transposition.**

La classification proposée par Finnegan (1989) a été complétée par Capy et al. (1997) qui ont subdivisé les deux classes en ordres, superfamilles et familles selon la fonction et la structure moléculaire des ET et la présence de motifs signatures. Les classifications récentes de Wicker et al. (2007) et Kapitonov et Jurka (2008) proposent de classer les ET en fonction de l'organisation de leurs domaines fonctionnels mais aussi de leur site d'insertion ou TSD (Target Site Duplication) dans le génome.

La classification proposée par Wicker et al. (2007) est la plus exhaustive des ET présents chez les eucaryotes mais elle exclut les ET des procaryotes qui sont nommés séquences d'insertion (IS). Dans la classification de Wicker et al. (2007) des ET des eucaryotes (Fig. 11), les deux classes principales constituent le premier niveau et les niveaux suivants sont déterminés sur la base des spécificités de la transposition, les structures d'éléments, les sites d'insertion et de séquences homologues. À un niveau inférieur, les superfamilles sont caractérisées par la structure des protéines ou des domaines non-codants et du site cible de duplication (TSD) des ET. Le TSD se caractérise par la duplication de la signature de la séquence adjacente à l'insertion des ET. Il y a environ 30 superfamilles connues actuellement, de nouvelles familles restent probablement encore à découvrir.

Classification		Structure	TSD	Code	Occurrence
Order	Superfamily				
<b>Class I (retrotransposons)</b>					
LTR	<i>Copia</i>	→ GAG AP INT RT RH →	4-6	RLC	P, M, F, O
	<i>Gypsy</i>	→ GAG AP RT RH INT →	4-6	RLG	P, M, F, O
	<i>Bel-Pao</i>	→ GAG AP RT RH INT →	4-6	RLB	M
	<i>Retrovirus</i>	→ GAG AP RT RH INT ENV →	4-6	RLR	M
	<i>ERV</i>	→ GAG AP RT RH INT ENV →	4-6	RLE	M
DIRS	<i>DIRS</i>	→ GAG AP RT RH YR →	0	RYD	P, M, F, O
	<i>Ngaro</i>	→ GAG AP RT RH YR → → →	0	RYN	M, F
	<i>VIPER</i>	→ GAG AP RT RH YR → → →	0	RYV	O
PLE	<i>Penelope</i>	← RT EN →	Variable	RPP	P, M, F, O
LINE	<i>R2</i>	← RT EN →	Variable	RIR	M
	<i>RTE</i>	← APE RT →	Variable	RIT	M
	<i>Jockey</i>	← ORF1 APE RT →	Variable	RIJ	M
	<i>L1</i>	← ORF1 APE RT →	Variable	RIL	P, M, F, O
	<i>I</i>	← ORF1 APE RT RH →	Variable	RII	P, M, F
SINE	<i>tRNA</i>	← →	Variable	RST	P, M, F
	<i>7SL</i>	← →	Variable	RSL	P, M, F
	<i>5S</i>	← →	Variable	RSS	M, O
<b>Class II (DNA transposons) - Subclass 1</b>					
TIR	<i>Tc1-Mariner</i>	→ Tase* →	TA	DTT	P, M, F, O
	<i>hAT</i>	→ Tase* →	8	DTA	P, M, F, O
	<i>Mutator</i>	→ Tase* →	9-11	DTM	P, M, F, O
	<i>Merlin</i>	→ Tase* →	8-9	DTE	M, O
	<i>Transib</i>	→ Tase* →	5	DTR	M, F
	<i>P</i>	→ Tase →	8	DTP	P, M
	<i>PiggyBac</i>	→ Tase →	TTAA	DTB	M, O
	<i>PIF-Harbinger</i>	→ Tase* ORF2 →	3	DTH	P, M, F, O
	<i>CACTA</i>	→ Tase ORF2 →	2-3	DTC	P, M, F
Crypton	<i>Crypton</i>	← YR →	0	DYC	F
<b>Class II (DNA transposons) - Subclass 2</b>					
Helitron	<i>Helitron</i>	← RPA Y2 HEL →	0	DHH	P, M, F
Maverick	<i>Maverick</i>	← C-INT ATP CYP POL B →	6	DMM	M, F, O

### Caractéristiques structurales



**Figure 11 : Classification des éléments transposables (d'après Wicker et al., 2007).**

Les ET eucaryotes de classe I et II sont divisés en sous-classes, ordres et superfamilles. La taille du fragment dupliqué après insertion ou TSD (Target Site Duplication) est caractéristique de la plupart des superfamilles. LTR: Répétitions terminales longues, TIR: Répétitions terminales inversées.

AP : protéinase aspartique  
C-INT: C-intégrase  
ENV: protéine d'enveloppe  
INT: intégrase  
POL B: ADN polymérase B  
RT: reverse-transcriptase  
YR : tyrosine recombinase

APE: endonuclease apurinique  
CYP: cystéine protéase  
GAG: protéine de la capsid  
ORF: cadre de lecture ouvert de fonction inconnue  
RH: Ribonucléase H  
Tase: transposase (\*avec un motif DDE)  
Y2YR: avec un motif YY

ATP: ATPase d'emballage  
EN: endonuclease  
HEL: hélicase  
RPA: protéine A de réplication

### Groupes de l'espèce

P: plantes ; M: métazoaires ; F: champignons ; O: autres

- Les ET de classe I sont classés en cinq ordres: LTR (Long Terminal Repeat), DIRS (Dictyotetium Intermediate Repeat Sequence), PLE (Penelope Like Element), LINE (Long Intersperced Nuclear Elements) et SINE (Short Intersperced Nuclear Elements) qui regroupent un total de 17 superfamilles qui sont caractérisées par l'organisation des gènes en domaines.

- Les ET de classe II sont répartis en quatre ordres: TIR, Crypton, Helitron et Maverick. L'ordre TIR comprend tous les ET présentant des répétitions terminales inversées (TIR: Terminal Inverted Repeats) à chaque extrémité, et il est le seul qui est divisé en neuf superfamilles. Ces superfamilles sont: *Tc1-mariner*, *hAT*, *Mutator*, *Merlin*, *Transib*, *P*, *PiggyBac*, *PIF Harbinger* et *CACTA*. Par exemple, la superfamille *Tc1-mariner* est composée de transposons qui partagent un gène homologue qui code la transposase, enzyme qui catalyse leur transposition. Une autre caractéristique de cette superfamille est que tous ces transposons possèdent comme TSD le dinucléotide TA.

Cette classification prend aussi en compte les ET non-autonomes dont on connaît mal le mécanisme de transposition tels que les MITE (Miniatures Inverse Transposable Elements), ils sont classés dans les superfamilles d'éléments autonomes de classe II. Selon Wicker et al. (2007), les MITE peuvent être classés en se basant sur l'homologie de séquences avec les ET autonomes permettant leur mobilisation en *trans*. Les MITE sont souvent définis comme des versions délétées non autonomes de transposons de classe II et sont mobilisables en *trans* par les éléments entiers dont ils dérivent. Selon la séquence et la taille du site cible dupliqué, les MITE peuvent être divisés en deux familles : les éléments *Tourist-like* (5'TAA-3') et les éléments *Stowaway-like* (5'-TA-3') qui appartiennent aux superfamilles *PIF-Harbinger* et *Tc1-mariner*, respectivement (Bureau et Wessler 1992; 1994; Feschotte et al., 2003 ; Zhang et al., 2004). Ils sont caractérisés par de courtes séquences sans capacité de codage, contiennent des TIR conservés, sont flanqués de TSD produits par l'insertion et proviennent probablement d'un sous-ensemble de transposons à ADN autonomes (Bureau et Wessler, 1992; Jiang et al., 2004). Il existe cependant des MITE *Tourist-like* et *Stowaway-like* qui présentent un cadre de lecture ouvert codant une transposase potentielle chez *A.thaliana* et *O. sativa* (Le et al., 2000 ; Turcotte et al., 2001). Les MITE ont été découverts chez les plantes, mais ils ont également été trouvés dans plusieurs génomes d'animaux (Feschotte et al., 2002 ; Gonzalé et Petrov, 2009).

La question de la classification n'est pas encore complètement résolue car de nombreux cas restent ambigus : l'abondance d'éléments incomplets tronqués ou fossiles (dont seule une partie très conservée a été identifiée) et l'afflux d'ET *de novo* découverts *in silico* grâce à l'analyse des génomes séquencés pourraient remettre en cause cette classification.

### **3.2. Cycle de vie des éléments transposables dans les génomes**

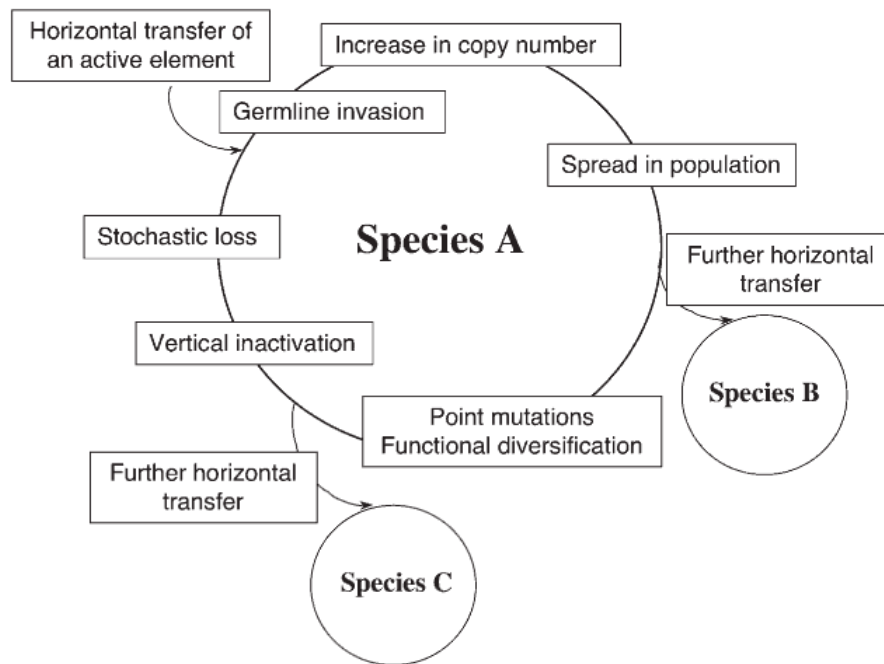
Selon le modèle de Kidwell et Lish (2001), le cycle de vie des ET se compose de trois étapes: i) l'invasion, ii) la maturation, iii) la sénescence). i) Un ET actif et autonome est inséré dans le génome probablement à la suite d'un transfert génétique horizontal (HGT), c'est dire entre espèces différentes (Robertson et Lampe, 1995 ; Silva et Kidwell, 2000 ; Jain et al., 2002 ; Diao et al., 2006). Le transfert de gènes est bien connu chez les bactéries (Jain et al., 2002) par contre les exemples de HGT sont moins nombreux entre les eucaryotes. Un exemple est donné par un transposon de type *Mutator* qui a été transmis par transfert horizontal entre deux lignées de Poacées (Diao et al., 2006). Après insertion, les ET sont capable de se répliquer dans le génome jusqu'à ce qu'il induise des effets délétères en modifiant la fonction des gènes et réduise ainsi le métabolisme de l'hôte ii) Lorsque les effets délétères sont observés, les processus de sélection entrent en jeu et induisent des mécanismes de régulation qui limitent la prolifération de l'ET. Ces mécanismes de régulation peuvent être, par exemple, la répression de la transcription par la méthylation de l'ADN. Ces mécanismes empêchent l'expression des gènes qui codent des enzymes de transposition (Machida et Machida, 1989; Zhang et al., 2001; Slotkin et Martienssen, 2007). iii) La dernière phase correspond à une diminution drastique du nombre de copies des ET originaux, jusqu'à ce qu'il disparaisse totalement du génome. Certains de ces processus pourraient être responsables de l'élimination totale des ET dans les génomes de certaines espèces qui n'ont pas subi une nouvelle invasion par HGT comme par exemple via des procaryotes intracellulaires obligatoires (Wright et Finnegan, 2001 ; Bordenstein et Reznikoff, 2005 ; Cordeaux, 2009).

La contre-sélection (Le Rouzic et Capy, 2005), la perte stochastique ou la dérive génétique peut conduire à l'élimination de certains ET du génome. Il a été suggéré que ces mécanismes sont responsables de la distribution sporadique de la famille de transposons *mariner* dans les populations de *Drosophila* (Lohe et al., 1995, Wang et al., 2005). Plusieurs phases d'invasion provenant d'autres organismes, comme des bactéries

ou des parasites sont également possibles à la suite d'événements de HGT. Bien qu'il soit difficile d'évaluer ces modes de transmission latéraux (Diao et al. 2006 ; Fortune et al., 2008), ils sont considérés comme ayant eu lieu à plusieurs reprises pour les ET (Robertson et Lampe, 1995 ; Silva et Kidwell, 2004 ; Roulin et al., 2008 ; Gilbert et al., 2010 ; Wallau et al., 2011) en particulier dans le milieu marin où ils peuvent être favorisés par le contact permanent avec les bactéries, virus et parasites (Sobecky et Hazen, 2009).

Comme indiqué précédemment, les éléments de la superfamille *Tc1/marinere* nécessitent pas de facteurs spécifiques de l'hôte pour transposer, ils ne sont donc pas dépendants d'un hôte spécifique. En effet, de nombreux cas de transfert horizontal entre différents organismes ont été proposés pour ces éléments. Les exemples rapportés dans la littérature concernent le transfert entre des crustacés marins (Casse et al., 2006), entre des insectes de différents ordres (Lampe et al., 2003 ; Robertson et Lampe, 1995), et même entre des organismes de phylums différents, aussi divergents que l'homme et un nématode parasite (Laha et al., 2007). Cependant, on ne sait pas comment ces éléments sont capables d'envahir de nouveaux génomes. Les vecteurs potentiels qui pourraient être impliqués dans ce transfert horizontal sont des parasites externes, tels que les acariens, qui semblent avoir été le vecteur des éléments *P* chez la drosophile (Houck et al., 1991), ou les parasites internes tels que les virus (Hartl et al., 1997a).





**Figure 12 : Évolution du cycle de vie des éléments *Tc1/mariner* chez les hôtes naturels. D'après Miskey et al. (2005).**

La figure 12 montre les principaux événements du cycle de vie (pour plus de détails, voir le texte). Le cycle a été proposé pour décrire l'évolution des éléments *mariner* (Hartl et al., 1997b), mais il est aussi probablement valable pour la famille *Tc1*. Le transfert horizontal de transposons actifs dans de nouvelles espèces peut avoir lieu avant ou après la diversification fonctionnelle.

### **3.3. Impacts des éléments transposables sur les génomes**

#### **3.3.1. Effets délétères des éléments transposables**

En raison de leur activité et de leur présence massive dans les génomes, les ET peuvent être considérés comme des séquences parasites qui provoquent des effets potentiellement délétères. La prolifération des ET augmente aussi la taille des génomes, notamment chez les plantes (Devos et al., 2002 ; Hawkins et al., 2006 ; Neumann et al., 2006 ; Piegu et al., 2006 ; Vitte et Panaud, 2005). Il semble évident que ces nombreux éléments doivent avoir affecté la structure du génome en facilitant des réarrangements chromosomiques et qu'ils perturbent l'expression des gènes (Bourque, 2009). Ainsi, des changements phénotypiques tels que l'apparition de pétalos panachés chez le mufler et le soja sont dus à l'insertion d'ET dans des gènes contrôlant la pigmentation des fleurs (Hashida et al., 2003 ; 2006 ; Takahashi et al., 2012). Comme nous l'avons vu, un effet de ce type a conduit à la découverte originale des

premiers éléments transposables Ds-Ac chez le maïs (McClintock, 1951) et celle de *mariner230* chez la drosophile (Hartl, 2001). Un autre exemple de perturbation génique, mais cette fois avec des conséquences néfastes, est celui de l'hémophilie, une maladie provoquée par l'insertion d'un retrotransposon L1 dans le gène du chromosome X humain qui code le facteur VIII de la coagulation du sang (Dombrosky et al., 1991). Toutefois, les ET n'ont pas seulement des effets délétères, sinon ils n'auraient pas été conservés dans les génomes pendant une si longue période. Se pose alors la question de leur conservation et des avantages que les ET peuvent apporter aux génomes qui les hébergent.

### **3.3.2. Éléments transposables sont-ils bénéfiques ?**

Nous pouvons penser que les ET sont conservés dans les génomes soit parce qu'ils ont des effets neutres ou parce qu'ils ont des effets positifs et confèrent certains avantages aux organismes. Il a été suggéré que les interactions entre les ET et le génome de l'hôte peuvent être considérées comme un continuum allant du parasitisme à l'interaction symbiotique (Kidwell et Lish, 2001). Les ET sont actuellement très étudiés car ils semblent de plus en plus jouer un rôle majeur dans l'évolution des génomes en générant de la diversité et en créant de nouveaux gènes. Les ET peuvent être des outils de régulation des gènes qui induisent la formation de nouveaux gènes : la forme extrême de ce phénomène est connue comme la "domestication" dans laquelle la séquence entière d'un ET est intégrée dans le génome et devient un acteur du fonctionnement de l'hôte. Il existe plusieurs exemples de ET dont une partie de leur séquence est devenue une séquence régulatrice de gènes tels que les promoteurs, des amplificateurs ou des séquences cis-régulatrices (Sinzelle et al., 2009 ; Bourque, 2009 ; Feschotte, 2008 ; Levin et Moran, 2011). Ils peuvent également conduire à la formation de gènes chimériques (une fusion entre une partie d'ET et certaines séquences de gènes) menant à une nouvelle protéine fonctionnelle qui peut être utile à l'organisme hôte. Le système immunitaire V(D)J des vertébrés en est un exemple, il dériverait d'un transposon *transib* (Kapitonov et Jurka, 2005). En effet, les similitudes fonctionnelles et structurales des protéines RAG1 et RAG2 impliquées dans la recombinaison V(D)J laissent penser que les gènes qui les codent sont des parties d'un ET actif possédant des TIR identiques aux signaux de cette recombinaison (Hiom et al., 1998).

Dans la plupart des cas, les mutations qui résultent de transpositions sont neutres, mais certains événements d'insertion ont des effets positifs pour les organismes et les

populations doivent avoir été sélectionnées au cours de l'évolution des génomes, comme c'est le cas pour les ET domestiqués.

### **3.4. Utilisation des éléments transposables**

#### **3.4.1. Outils de mutagenèse et de transgénèse**

Certains ET sont utilisés *in vitro* et *in vivo* comme outils de mutagenèse et de transgénèse aussi bien chez les cellules animales que végétales (Dinesh-Kumar, 1995). Chez les plantes, le contrôle de la transposition des éléments *Ac/Ds* ont permis le développement de la mutagenèse dirigée chez le maïs, mais aussi le tabac, la tomate et *Arabidopsis thaliana* (Sundaresan, 1996). La découverte de transposons de type *mariner* potentiellement actifs permet d'envisager la mise au point d'un système de transgénèse ubiquiste. Les premiers essais de transgénèse ont été réalisés chez *Drosophila melanogaster* à l'aide de l'élément *P* (Rubin et Spradling, 1982). Chez la cione *Ciona intestinalis*, la transposase du transposon *Minos* possède une activité complète puisqu'elle est capable d'exciser de l'ADN et de transposer (Sasakura et al., 2003).

La thérapie génique est une stratégie très prometteuse dans le traitement de plusieurs maladies génétiques héréditaires ou acquises. Divers types de vecteur ont été utilisés pour délivrer des séquences d'ADN "thérapeutique" aux cellules affectées. Les vecteurs viraux ont montré des effets génotoxiques et des complications immunologiques alors que les vecteurs non viraux sont moins efficaces dans l'intégration au niveau du génome. Une solution alternative consiste à utiliser les ET comme vecteurs, ces éléments ayant la capacité de s'intégrer facilement dans le génome et de permettre l'introduction d'un gène d'intérêt et son expression (Ivics and Izsvák, 2006).

Les transposons ont trois caractéristiques qui les rendent attractifs (Nesmelova et Hackett, 2010) : i) La transposition est précise lorsque la séquence génique qui est mobilisée pour l'intégration est bien définie. ii) La transposition assure l'intégration de transposons simples. iii) Contrairement aux virus qui sont souvent couverts de protéines hautement immunogènes, les transposons comprennent uniquement des fragments d'ADN et ils évitent ainsi des réactions immunitaires et d'autres mécanismes de défense que les cellules utilisent pour empêcher l'intégration d'ADN étranger. Par conséquent, les ET présentent un intérêt substantiel dans l'amélioration de l'efficacité transpositionnelle.

Quelques transposons sont en cours de développement pour être utilisés chez les vertébrés : *Sleeping Beauty* (Ivics et al., 1997 ; Hackett et al., 2004), *Frog Prince* (Miskey et al., 2003), et *Tol2* (Balciunas et al., 2006). La reconstruction moléculaire de *Sleeping Beauty*, qui est un issu de poisson, marque une étape importante dans l'application de la thérapie génique fondée sur la transposition (Izsvak et Ivics, 2004; Ivics et Izsvák, 2006).

### **3.4.2. Marqueurs de polymorphisme**

L'activité de transposition des ET crée de la variabilité génétique dans toutes les espèces. Il existe un polymorphisme d'insertion qui est utilisé comme outil de marquage moléculaire pour évaluer la diversité génétique. Les ET utilisables doivent présenter un grand nombre de copies, une potentialité de polymorphisme élevée, et une stabilité des insertions. Les rétrotransposons à LTR et les SINE répondent à ces critères et sont préférentiellement utilisés.

La détection des polymorphismes d'insertion nécessite des profils d'insertion comprenant l'ensemble ou une fraction significative des fragments de bordure des insertions d'un ET donnés dans un génome. La technique de SSAP (*Sequence-Specific Amplified Polymorphisms*) basée sur le rétrotransposon *BARE-1* a été appliquée dans des cartes génétiques chez l'orge (Waugh et al., 1997), chez l'avoine (Yu et Wise, 2000) et celle basée sur le rétrotransposon *PDR1* chez le pois (Ellis et al., 1998). Ces études ont montré que les marqueurs moléculaires basés sur le polymorphisme d'insertion des ET sont de puissants outils d'analyse de la biodiversité car ils présentent souvent un polymorphisme plus important que d'autres marqueurs moléculaires classiques. Par exemple, les intégrations récentes de l'élément *Alu* dans le génome humain ont généré des variantes qui ont une grande utilité dans l'étude des populations humaines. Ces variantes peuvent être utilisées comme des marqueurs dans les analyses de paternité et en médecine légale (Kass et al., 2007).

### **3.5. Transposons *mariner***

Le premier élément *mariner* a été découvert chez la drosophile *Drosophila mauritiana* par l'analyse d'une mutation instable dans l'allèle du gène *white* qui se traduit par un changement de coloration au niveau de l'œil. Cet élément appelé *Dmmar* (pour *Drosophila mauritiana mariner*) est plus connu sous le nom de *Mos1* (Jacobson et al., 1986). Les éléments apparentés à *Mos1* sont appelés *mariner*-like elements (MLE). Les MLEs sont largement répandus et représentés dans divers taxons tels que : les plantes, champignons, ciliés, rotifères (Arkhipova et al., 2005), insectes (Robertson, 1993), nématodes, poissons, mammifères, crustacés (Casse et al., 2006 ; Plasterk et al., 1999 ; Robertson, 2002) ou encore chez le protozoaire *Trichomonas vaginalis* (Silva et al., 2005).

#### **3.5.1. MLE chez les animaux**

Les MLE d'origine animale sont répartis en neuf sous-familles dont les noms dépendent des espèces dans lesquelles les premiers éléments ont été découverts : *mauritiana*, *cecropia*, *mellifera*, *capitata*, *irritans* et *elegans*, *briggsae*, *rosa* et *lineata* (Robertson, 1993, 1997, 2002 ; Arkhipova et Meselson, 2000; Shao et Tu, 2001; Feschotte et Wessler, 2002). Récemment, une nouvelle sous-famille *marmoratus* a été mise en évidence par Bui et al. (2007, 2008) chez les crabes marins. Cette sous-famille compte un élément *mariner* potentiellement actif *Pacmmar1* chez le crabe *Pachygrapsus marmoratus* (Delaurière et al., 2009).

Comme dans les autres familles d'éléments, la majorité des éléments *mariner* sont inactifs à cause de l'apparition de mutations au niveau de leur séquence. Chez les animaux, seuls deux éléments *mariner* naturels sont actifs et autonomes, *Mos1*, isolé chez la drosophile *Drosophila mauritiana* (Jacobson et al., 1986) et *Famar1* isolé chez le perce-oreille *Forficula auricularia* (Barry et al., 2004). Toutefois les éléments identifiés chez l'Homme *Himar1* et *Hsmar1* et chez les poissons *Sleeping beauty* ont été réactivés par reconstruction de la séquence ancestrale (Lampe et al., 1996 ; Miskey et al., 2007).

### **3.5.2. MLE chez les végétaux**

Les études concernant les MLE de plantes sont peu nombreuses. La première étude a été menée par Jarvick et Lark en 1998. Ils ont isolé dans le génome du soja un élément qu'ils ont nommé *Soymar1* (*Soybean mariner* élément 1) présentant des similitudes avec l'élément *mariner*. L'alignement des séquences protéiques de transposases montre que près de la moitié de cette transposase diffère des transposases *mariner* présentes chez la drosophile (*Mos1*), chez l'homme (*Himar1*) ainsi que chez les planaires. Toutefois, l'élément *Soymar1* possède la triade catalytique DDxD typique des MLE et nécessaire à la transposition mais la distance x entre les deux derniers acides aspartiques est de 39 acides aminés au lieu de 34 chez les animaux (Jarvick et Lark, 1998). Cette triade DD39D est aussi caractéristique de l'élément *Osmar1* (*Oryza sativa mariner* élément 1) identifié chez le riz par Tarchini et al. (2000). Ces deux MLE végétaux partagent 67 % d'homologie. Une recherche dans les bases de données a permis d'identifier deux autres transposons complets de type MLE dans le génome du riz, *Osmar2* et *Osmar2b* (Feschotte et Wessler, 2002). Ces auteurs ont mis en évidence l'ubiquité des MLE chez les plantes puisqu'ils ont amplifié des fragments de gènes codant une transposase de MLE chez 36 espèces d'angiospermes. Puis d'autres MLE ont été découverts chez d'autres plantes dont un élément complet et potentiellement actif *Vulmar1* chez la betterave *Beta vulgaris* (*Beta vulgaris mariner* élément 1) (Jacobs et al., 2004), cinq MLE ont été découverts chez le pois, dont un élément complet potentiellement actif *Psmar1* (*Pisum sativum mariner* élément 1) (Macas et al., 2005), 79 gènes de la transposase MLE ont été isolés chez 63 espèces de bambou (Zhou et al., 2011) dont la grande majorité contient de multiples mutations inactivant ces éléments. Tous les MLE complets de plantes possèdent la triade catalytique DD39D.

Les séquences des ITR des MLE de plantes montrent une forte homologie avec les séquences des ITR des MITE *Stowaway*. Ceci suggère que ces MITE peuvent être mobilisées par des transposases de MLE fonctionnels.

### **3.5.3. Éléments transposables chez les microalgues**

Les données disponibles dans la littérature sur les ET chez les différents groupes de microalgues sont rares. Chez la microalgue verte *Chlamydomonas reinhardtii*, deux ET *Gulliver* et *TOC1* ont été découverts par Ferris (1989) et Day et al. (1988)

respectivement. *Gulliver* est un transposon de classe II qui ressemble à l'élément *Ac* du maïs (Graham et al., 1995). L'élément TOC1 est un rétrotransposon à LTR inséré dans un intron du gène OEE1 (Oxygen-evolving enhancer 1) du mutant FUD44. Des éléments nommés *Pioneer1* et *Tcr3* ont été identifiés par Graham et al. (1995) et Wang et al. (1998), ces éléments insérés dans les gènes de la nitrate réductase seraient des transposons ayant perdu une extrémité.

Le génome de quatre diatomées marines *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, *Fragilariopsis cylindrus* et *Pseudonitzschia multiseries* ont été séquencés. Le séquençage complet de ces génomes permet de rechercher *in silico* des ET et de comparer ces données aux recherches expérimentales au laboratoire. Bien que la taille de ces génomes soit relativement faible 32,4 Mpb pour *T. pseudonana* et 27,4 Mpb pour *P. tricornutum* (Armbrust et al., 2004, Bowler et al., 2008 ; Vardi et al., 2008), les ET semblent être les responsables majeurs des réarrangements génomiques et de la dynamique de ces génomes. *P. tricornutum* contient une proportion de 6,4 % d'ET, soit trois fois plus que *T. pseudonana* qui ne possède que 1,9 % d'ET. Les rétrotransposons à LTR-RT de type *Copia-like* représentent 90 % des ET de *P. tricornutum* (soit 5,8 % du génome) et 58 % chez *T. pseudonana*, (soit ~ 1 % du génome) (Armbrust et al., 2004; Bowler et al., 2008; Maumus et al., 2009 ; Tirichine et Bowler, 2011 ; Vardi et al., 2008). Les autres rétrotransposons appartiennent aux familles *Ty3/Gypsy*, *Ty1/Copia* et *DIRS-like*.

Dans notre laboratoire, des transposons de type *mariner* ont été recherchés expérimentalement chez 15 espèces de microalgues marines (Hermann, 2011). La présence de MLE a été révélée par la technique de PCR avec les amorces dégénérées définies par Feschotte et Wessler (2002). Des fragments de 380 pb identifiés positivement comme MLE ont été amplifiés chez 10 diatomées marines dont *Amphora acutiuscula*, provenant de la côte sud-est vietnamienne. L'absence de codon stop dans la région amplifiée du gène de la transposase suggère l'existence potentielle d'un élément complet de type *mariner* chez cette espèce (Hermann, 2011). Chez *P. tricornutum*, Hermann (2011) a mis en évidence plusieurs copies de MLE, par contre ces éléments semblent absents chez *T. pseudonana*.

Récemment, un transposon *Tc1-mariner* complet nommé *Costa* a été identifié *in silico* dans le génome de *Pseudonitzschia multiseries*, une diatomée produisant des peptides toxiques (<http://www.ncbi.nlm.nih.gov/nucleotide/JX013907> ).

### 3.5.4. Structure des transposons MLE

Les MLE sont les plus petits transposons ADN trouvés chez les eucaryotes, leur taille étant comprise entre 1,3 Kpb chez les animaux et 7,072 Kpb chez les plantes (Plasterk et al., 1999; Leroy et al., 2003). Ces transposons contiennent un seul gène qui code la transposase qui est l'enzyme catalysant la transposition. De part et d'autre du gène de la transposase se trouvent des régions terminales non traduites ou UTR (UnTranslated Region) dont la séquence est variable. Les MLE se terminent par des répétitions terminales inversées (ITR) qui sont mal conservées, dont la taille varie de 20 à 40 pb (Jarvik et Lark, 1998 ; Brillet et al., 2007 ; Feschotte et al., 2003) (Fig. 13-A).

Le gène de la transposase a une structure simple avec un cadre de lecture unique. Ce gène ne possède pas d'introns chez les MLE présents chez les animaux (Robertson, 2002) mais peuvent en contenir chez les MLE de plantes (Feschotte et Wessler, 2002 ; Feschotte et al., 2003 ; Zhou et al., 2011).

Les ITR des MLE ont une taille d'environ 20 à 40 bp, que ce soit chez les animaux ou les végétaux (Jarvik et Lark, 1998 ; Brillet et al., 2007; Jacobs et al., 2004; Macas et al., 2005 ; Feschotte et al., 2003). Ils possèdent un motif palindromique où chacune des deux parties du palindrome posséderait un site de liaison pour la transposase et un motif de miroir dont le rôle serait de stabiliser les interactions protéines-ITR (Bigot et al., 2005).

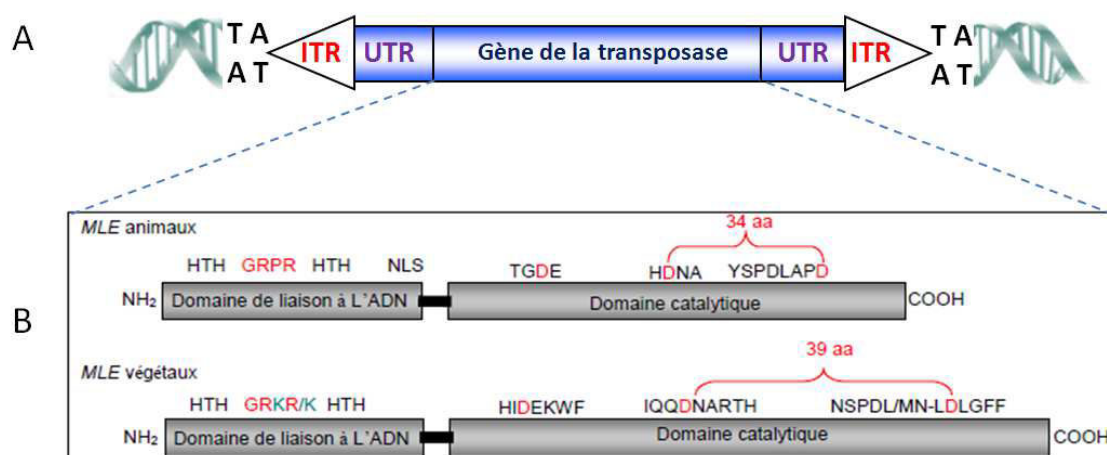


Figure 13 : Structure des éléments *mariner*.



(A) Représentation schématisée d'un transposon MLE. Les répétitions inversées terminales contiennent le site de liaison pour la transposase. Un seul gène codant la transposase est encadré par des séquences UTR et ITR. Le site cible de dinucléotide TA est présent à chaque extrémité du transposon, il est dupliqué au niveau du site d'intégration lors d'une transposition.

(B) Représentation de la structure des transposases des MLE animaux et végétaux. Le domaine de liaison à l'ADN se trouve à l'extrémité N-terminale. Dans ce domaine, deux motifs hélice-tour-hélice (HTH) sont séparés par un motif commençant par GRxR conservé au sein de la famille MLE et sont suivis du signal de localisation nucléaire (NLS). Le domaine catalytique à l'extrémité C-terminale est responsable de la catalyse du clivage. La triade d'acides aminés aspartiques DDxD est une signature caractéristique des transposases MLE. La distance de 34 acides aminés et celle de 39 sont très conservées pour les MLE animaux et végétaux, respectivement.

Les UTR des MLE sont des régions non traduites situées en 5' et 3' du cadre de lecture codant la transposase. Ces régions peuvent atteindre plusieurs milliers de nucléotides et sont la principale cause de la variation de taille observée au sein des TLE (Sinzelle et al., 2005). Leur rôle est peu connu. Ils pourraient posséder les informations nécessaires à la transcription et la traduction de la transposase. Ainsi, ils pourraient permettre la fixation des facteurs de transcription (Leroy et al., 2003 ; Palomeque et al., 2006). Les ITR seuls sont suffisants pour que la transposase effectue la transposition (Augé-Gouillou et al., 2001a ; Feschotte et al., 2003), cependant la présence d'au moins une partie des UTR permettrait une transposition plus efficace (Jeyaprakash et Hoy, 1995; Miskey et al., 2007 ; Tosi et Beverley, 2000).

### **3.5.5. Structure protéique des transposases de MLE**

Chez les animaux, cette protéine contient plusieurs motifs hautement conservés dont WVPHE et YSPDLAP qui sont les signatures principales des transposases de MLE (Robertson, 1993; Plaster et al., 1999). Les transposases ont une taille comprise entre 330 et 360 acides aminés chez les animaux et de 412 à 520 acides aminés chez les végétaux. Elles sont constituées de deux domaines fonctionnels : un domaine de liaison à l'ADN situé dans la région N-terminale et un domaine catalytique situé en région C-terminale (Fig. 13-B). Dans la région N-terminale, les transposases possèdent deux motifs hélice-tour-hélice (HTH) nécessaires pour la reconnaissance et la liaison aux TIRs et un ou plusieurs signaux de localisation nucléaire (NLS). Le domaine C-terminal héberge la triade catalytique constituée par trois acides aspartiques (DDD) (Fig. 13-B). Les deux premiers aspartates sont séparés par 92 acides aminés, tandis que la

distance entre le second et le troisième résidu est variable, entre 31 et 39 acides aminés selon les familles (Brillet et al., 2007).

### ***Domaine de liaison à l'ADN***

Le domaine de liaison à l'ADN est responsable de la reconnaissance des séquences spécifiques des ITR du transposon. Ce domaine a été déterminé par des analyses de prédiction de structure secondaire (Petrokovski et Henikoff, 1997) et de cristallographie des transposases de l'élément *Mos1* (Richardson et al., 2004) et par des études moléculaires d'interactions entre la transposase et les TIR. Ce domaine est généralement situé dans les 150 premiers acides aminés de la transposase et est constitué de cinq hélices. Les deux dernières hélices forment un motif hélice-tour-hélice (HTH) de fixation à l'ADN de type *homeo-like*, précédé de la séquence conservée GRPR (Plasterk et al., 1999 ; Vos et Plasterk, 1994). Selon les auteurs, les deuxième et troisième hélices forment également un motif HTH. Dans le cas de *Mos1*, l'hypothèse d'un seul motif est proposée par Augé-Gouillou et al. (2001b), alors que l'hypothèse de deux HTH est décrite par Richardson et al. (2006).

### ***Signal de localisation nucléaire***

La région N-terminale contient un ou plusieurs signaux de localisation nucléaire (NLS) permettant le transport de la transposase dans le noyau de la cellule (Brillet et al., 2007). Pour la plupart des éléments, ce NLS a été défini *in silico* par analyse bioinformatique et n'est que putatif. En ce qui concerne les MLE de plantes, seuls les travaux de Yang et al. (2006) sur la transposition de *Osmar5* en système hétérologue ont permis de définir le domaine de liaison à l'ADN et les domaines hélice-tour-hélice ainsi que le mécanisme de transposition des MLE de plantes.

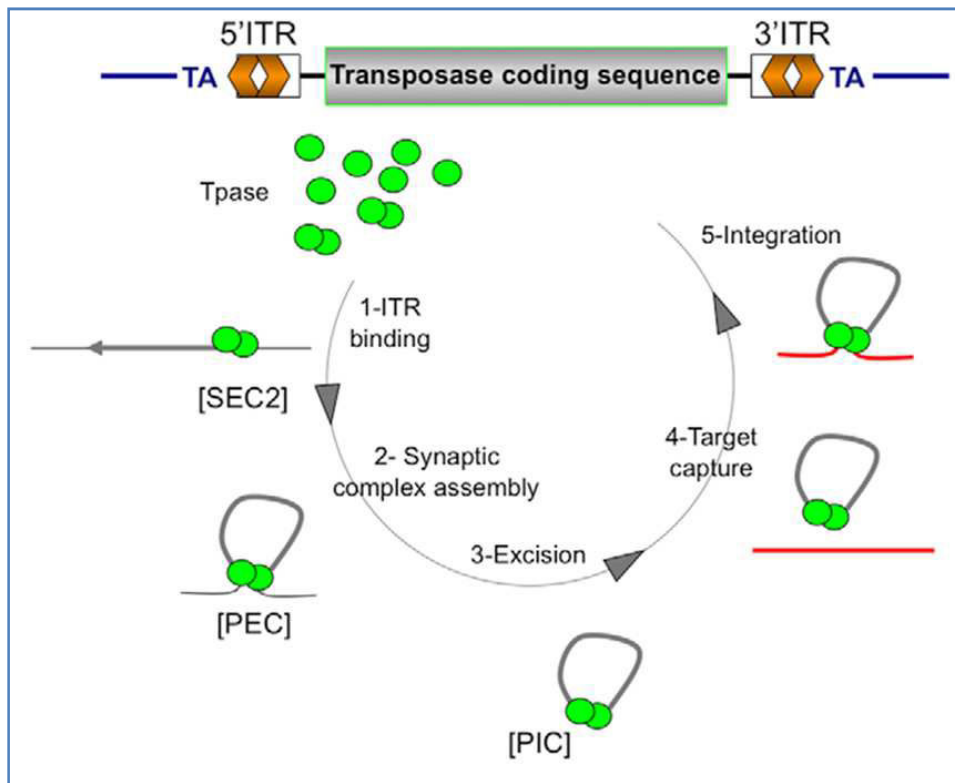
### ***Domaine catalytique des transposases***

Le domaine catalytique des transposases de type MLE est caractérisé par la triade catalytique DDxD. Le dernier D semble essentiel à l'activité de la transposase *Mos1*, son remplacement par un acide glutamique E (DD34E au lieu de DD34D) entraîne une perte de la fonctionnalité protéique (Lohe et al., 1997). Cependant dans le cas de *Hsmar1* le remplacement de l'acide aspartique (D) par l'asparagine (N) ne modifie pas l'activité *in vitro* de la transposase (Liu, et al., 2007). Pour assurer son rôle, la triade fixe un à deux cations divalents comme le magnésium ou le manganèse (Mizuuchi, 1997; Richardson et al., 2009). Les trois acides aspartiques sont inscrits dans les motifs conservés TGDE, HDNA et YSPDLAPxD des éléments isolés chez les animaux. Ce

dernier motif a été décrit comme l'un des deux motifs signatures des MLE, avec le motif WVPHEL (Robertson, 1993), bien qu'aucun rôle ne lui ait encore été attribué. Les éléments isolés chez les végétaux présentent une triade catalytique légèrement plus longue. En effet, 39 acides aminés séparent les deux derniers acides aspartiques (Jarvick et Lark, 1998; Tarchini et al., 2000; Jacobs et al., 2004; Macas et al., 2005). Les motifs conservés ont été identifiés comme étant HIDEKWF et IQQDNARTH. Le troisième résidu aspartate de la triade catalytique semble apparaître dans un motif NSPDL/MNxLDLGFF (Fig. 13-B). Toutefois, l'étude de la transposition de *Osmar5* en système hétérologue montre qu'une mutation du deuxième résidu aspartate de la triade catalytique induit une nette diminution de la transposition n'allant pas jusqu'à inhiber totalement l'activité de la transposase (Yang et al., 2006). Les auteurs suggèrent que le motif DD39D, défini par la conservation des motifs, pourrait ne pas refléter le véritable centre réactionnel de la transposition des MLE chez les plantes. Ainsi, les motifs de la triade catalytique restent à définir par l'étude de la transposition d'autres MLE de plantes.

### **3.5.6. Mécanisme de transposition des MLE**

L'étude du mécanisme de transposition des éléments *mariner* a été réalisée essentiellement sur *Mos1* (Dawson et Finnegan, 2003; Augé-Gouillou et al., 2005 ; Bigot et al., 2005 ; Richardson et al., 2009 ; Carpentier et al., 2011). La transposition des éléments *mariner* implique une série de cinq étapes: (1) liaison spécifique de la transposase aux ITR, (2) clivage des deux brins d'ADN à chaque extrémité du transposon, (3) liaison des deux extrémités du transposon par des interactions transposase-transposase, (4) capture d'une nouvelle cible qui est un dinucléotide TA et (5) insertion de l'élément dans ce locus.



**Figure 14 : Mécanisme de transposition de *mariner*.** D'après Jaillet et al. (2012).

Un élément *mariner* est représenté en haut avec ses principales composantes : les ITR (ITR-5' et ITR-3', en orange) et le dinucléotide TA flanquant l'élément. Il contient une séquence codant pour la transposase (en gris). SEC2: single-end complex, PEC: paired-end complex, PIC: pre-integration complex.

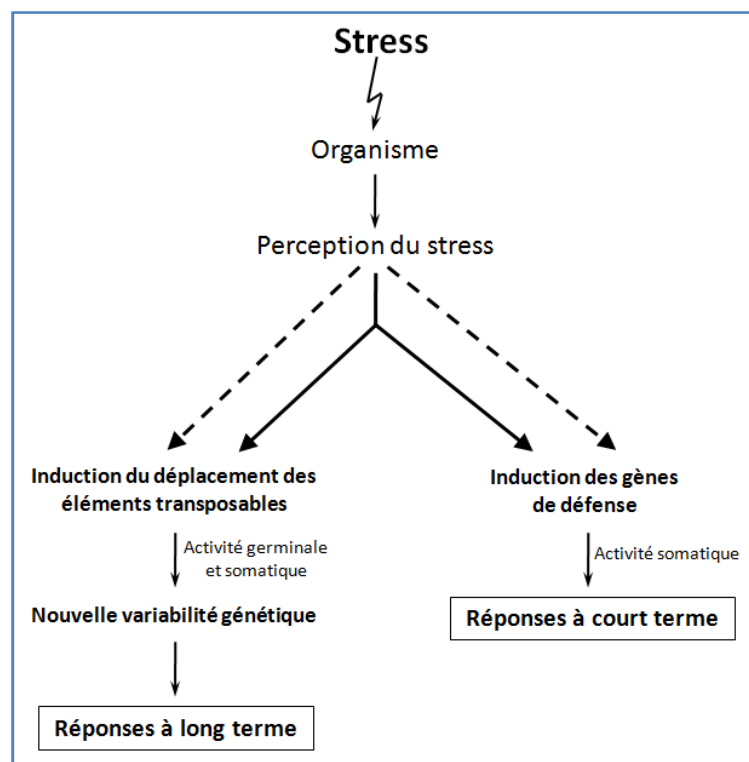
### ***Facteurs influençant l'efficacité de transposition***

La fixation spécifique de la transposase sur les ITR est dépendante des cations  $Mg^{2+}$ ,  $Mn^{2+}$  ou  $Ca^{2+}$  (Augé-Gouillou et al., 2001b ; 2005 ; Richardson et al., 2006). Le site catalytique de Mos1 renferme soit un ion  $Mg^{2+}$  soit deux ions  $Mn^{2+}$ , même en l'absence de substrat ADN (Richardson et al., 2006). La plupart des transposases de type DDE chez les procaryotes clivent, aux extrémités du transposon, le premier brin d'ADN par hydrolyse puis le second par une trans-estérification. A chaque extrémité, les deux brins se raboutent formant une structure intermédiaire en épingle à cheveux (DNA-hairpin) (Bischerour et al., 2009). Lors de la transposition des éléments *mariner*, la structure intermédiaire en épingle est absente (Richardson et al., 2006). Un premier brin d'ADN est coupé en 5' trois bases à l'intérieur de l'élément, puis l'extrémité 3' est coupée précisément à l'extrémité du transposon. Augé-Gouillou et al. (2005) ont montré que la transposase est active sous forme de dimère, la dimérisation se produisant avant la liaison aux ITR (Fig. 14). Il se forme ainsi un complexe composé de deux transposases (dimère) et de l'ITR à chaque extrémité d'un *mariner* excisé. Puis, les extrémités du transposon se rejoignent formant un complexe synaptique (complexe ADN-protéines) avec un dimère ou un tétramère de transposases. Le transposon est ensuite inséré au niveau du site TA.

La teneur en GC ainsi que la longueur de la partie interne du transposon *Mos1* ont une importance cruciale sur l'efficacité de la transposition (Sinelle et al., 2008). La transposition de *mariner* dépend aussi de la température. Les températures élevées inhibent l'auto-intégration des produits d'excision dans le génome, ce phénomène pourrait être en rapport avec les propriétés enzymatiques de la transposase (Sinelle et al., 2008).

#### **4. Stress et éléments transposables**

McClintock (1984) a été la première à suggérer que les éléments transposables pouvaient être activés en réponse à des stress, les génomes nécessitant des réarrangements pour survivre. Selon Li et al. (1999), les ET feraient aussi partie de la cascade de réactions en réponse aux stress et pourraient avoir un rôle de signal d'alarme. L'application de stress à des organismes dans des conditions contrôlées et l'étude de populations vivant dans des environnements différents permettent de préciser l'impact des stress sur les ET (Capy et al., 2000 ; Okamoto et Hirochika, 2001 ; Lukens et Zhan, 2007). Il ressort que seulement quelques éléments connus sont activés en conditions stressantes et que les réponses diffèrent selon les souches ou les populations. De plus, ce sont essentiellement les rétrotransposons qui sont activés. Toutefois, des transposons de la famille *mariner* sont activés dans des populations de *D. simulans* vivant à différentes latitudes (Giraud et Capy, 1996). Une diminution de température de 29 à 25 °C active aussi les transposons *P* et *hobo* chez *D. melanogaster* (Kidwell, 1977 ; Blackmann et al., 1987).



**Figure 15 : Induction du déplacement des éléments transposables après un stress (D'après Capy et al., 2000).**

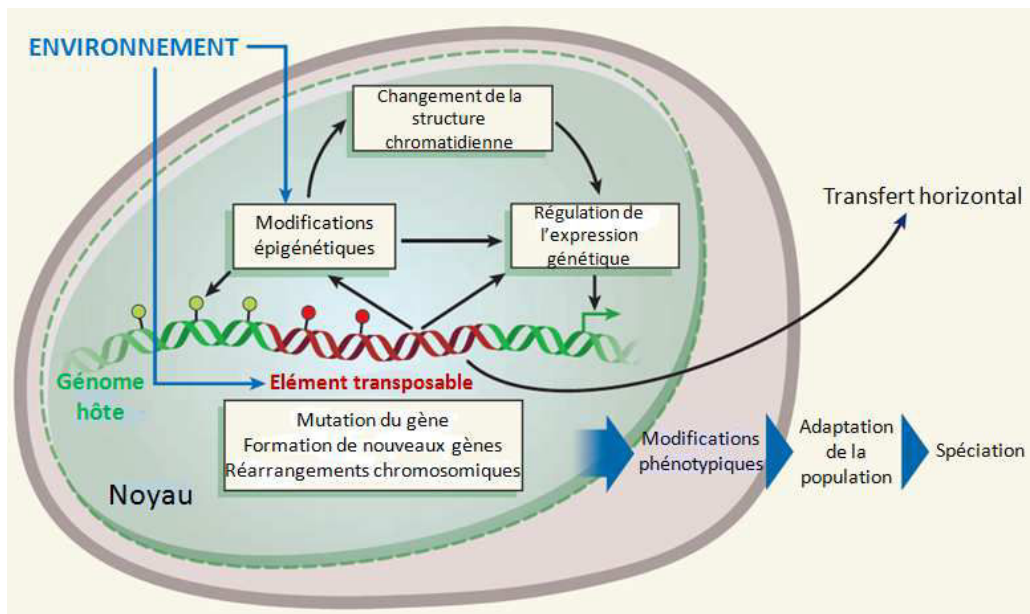
Les effets indépendants et communs sur les gènes de défense et les éléments transposables sont indiqués par des traits discontinus et continus, respectivement.

Chez les plantes, plusieurs cas d'activation d'ET ont été rapportés : celle du transposon *Tam3* est fréquemment activée à 15 °C et très rarement à 25 °C chez le muflier *Antirrhinum majus* (Harrison et Fincham, 1964 ; Carpenter et al., 1987), *Mutator* est le transposon le plus mutagène chez le maïs (Diao et Lisch, 2006) et une courte séquence MITE nommée *mPing* (superfamille *PIF/Harbinger*) s'est révélée active dans une lignée cellulaire d'*Oryza indica* (Jiang et al., 2003 ; Nakazaki et al., 2003). Chez le tabac et la tomate, un abaissement de température comme une attaque fongique induit la transcription d'un rétrotransposon *Tnt1* (Mhiri et al., 1997 ; Melayah et al., 2001 ; Grandbastien et al., 2005). Chez la diatomée modèle *P. tricornutum*, différents stress abiotiques tels qu'une carence en fer, en nitrates ou en CO<sub>2</sub> ou un excès de fer, de silice ou la présence d'un composé toxique (aldéhyde décadiénal) ont été appliqués afin de déceler d'éventuelles activations d'ET. Certains stress ont activé deux rétrotransposons de la superfamille *Ty1-copia* (Maumus et al., 2009), *Blackbeard* est celui qui est le plus fortement activé chez *P. tricornutum* carencé en nitrates et *Surcouf* est fortement exprimé en présence de l'aldéhyde. L'hypométhylation de l'élément *Blackbeard* a été mise en évidence chez les cellules carencées en nitrates (Maumus et al., 2009). Par ailleurs, une forte densité cellulaire active aussi *Blackbeard* chez *P. tricornutum* (Oliver et al., 2010).

Un modèle de l'impact des stress sur l'activité des ET à court terme et long terme a été proposé par Capy et al. (2000) (Fig. 15). Un stress peut être appliqué localement ou à un organisme entier et induire l'expression de gènes de défense permettant l'acclimatation de l'organisme à la contrainte (réponse à court terme). De la même manière, les ET peuvent être transcrits sous l'effet du stress et se déplacer dans le génome entraînant un remaniement qui affectera ou non l'expression des gènes selon leur lieu d'insertion. Cette variabilité génétique sera transmise aux nouvelles générations (effet à long terme). Parallèlement à ces inductions indépendantes, l'expression des ET seraient dans la majorité des cas induite indirectement. En condition de stress, des gènes de défense sont activés par des activateurs de transcription. Or, dans la région régulatrice des ET se trouvent des sites de fixation pour ces activateurs de transcription. Il n'est donc pas étonnant que les stress induisent aussi la transcription des ET ayant ces sites de fixation. Un stress biotique ou abiotique peut réactiver un rétrotransposon *Tnt1* chez le tabac, la région *cis* de cet élément contient des régions homologues aux boîtes G et H présentes dans les régions régulatrices de plusieurs gènes de défense chez les plantes (Mhiri et

al., 1997 ; Grandbastien, 1998 ; Beguiristain et al., 2001). De même, certains ET chez *Drosophila melanogaster* qui répondent au stress thermique contiennent également des motifs similaires aux promoteurs de gènes impliqués dans le choc thermique (Lerman et al., 2003). Le déplacement des ET dans le génome peut avoir des effets délétères dans le cas de mutations mais dans la plupart des cas, les remaniements génétiques sont localisés dans l'hétérochromatine ou subissent un contrôle épigénétique. Par ailleurs, le stress peut inhiber des mécanismes qui répriment habituellement les ET.

Chez *Drosophila simulans*, la température influe sur le taux de transposition du rétrotransposon nommé 412 et sur son nombre de copies dans les populations sauvages (Vieira et al., 1998). Une fois réactivés, ces éléments peuvent ensuite davantage modifier le génome de manière aléatoire pour répondre au stress. Cette variabilité transmise à la descendance favoriserait une meilleure adaptation des organismes aux changements environnementaux. Les ET sont de plus en plus considérés comme les facteurs majeurs de l'évolution et de l'adaptation des génomes, permettant aux populations de s'adapter aux changements environnementaux (Kidwell et Lish, 1997 ; Biémont et al., 2006 ; Tenaillon et al., 2010), comme décrit dans la figure 16.



**Figure 16 : Impact des facteurs de l'environnement sur les éléments transposables et le génome (Biémont et Vieira, 2006).**



Le contrôle épigénétique des ET inhibe la transposition ou rend les ET "silencieux". Les principaux mécanismes connus à ce jour sont la méthylation de l'ADN, la dégradation des transcrits d'ET et les modifications post-traductionnelles des histones (Rigal et Mathieu, 2011).

Woodcock et al. (1997) ont suggéré que la méthylation des transposons induite par des facteurs environnementaux jouait un rôle important dans leur inactivation. Ce type de répression de l'expression des ET se déroule dans le noyau et est transmis aux gamètes après la méiose (Vaucheret et Fagard, 2001). Plusieurs études ont mis en évidence le contrôle de l'activité des ET par la méthylation de l'ADN chez différents eucaryotes (Chandler et Walbot, 1986 ; Banks et al., 1988; Lippman et al., 2003). L'addition de groupements méthyles se produit essentiellement au niveau des cytosines de l'ADN des eucaryotes (Feng et al., 2010 ; Zemach et al., 2010).

Chez les plantes et les champignons, ces méthylations sont catalysées par des méthyltransférases (MET) guidées par des petites séquences d'ARN de 24 nucléotides appelés siRNA (small interfering RNA) (Cao et al., 2003; Matzke et al., 2009). Les siRNA peuvent provenir de transcrits d'ET coupés par une RNase III connue sous le nom de type Dicer ; ils aussi guident des protéines Argonautes qui clivent et bloquent alors la traduction d'ARN messagers cibles (Li et al., 2012). Chez les métazoaires, ce sont des Piwi ARN Interférant (piRNA) qui sont impliqués dans le "silencing" des ET. Les micro ARN (miRNA) d'une taille de 21 nucléotides résultent également de clivage d'ARN catalysés par Dicer ou des enzymes apparentées ; ils s'hybrident par appariements de base incomplets à plusieurs sites des ARNm cibles. Les miRNA répriment ainsi la traduction. Toutefois, les miRNAs présentent des caractéristiques différentes des siRNAs comme leur structure originelle simple brin qui s'organise en épingle à cheveux (hairpin). Les TE pourraient être une source de miRNAs impliqués dans la régulation des gènes chez les plantes (Li et al., 2012). La régulation des ET dépend aussi de diverses modifications post-traductionnelles des histones, la répression de la transcription de certains ET est contrôlée par la méthylation de celles-ci. Chez le riz, des mutations au niveau des gènes histone-H3K9-méthyltransférases entraînent l'activation d'ET (Ding et al., 2007). Le maintien de l'inactivation des ET chez *A. thaliana* semble fortement dépendre des histones déacétylées (Lippman et al., 2003) ; toutefois, compte tenu des fortes interactions entre la méthylation de l'ADN et les

modifications des histones, il est difficile d'évaluer la contribution de ces deux mécanismes sur l'inactivation des ET.

L'inhibition post-transcriptionnelle se déroule dans le cytosol où les ARN messagers sont dégradés, ce type d'inhibition n'est pas transmis à la génération suivante. Chez le mutant *mut6* de la microalgue verte *Chlamydomonas reinhardtii*, la quantité de transcrits du rétrotransposon LTR *TOC1* est trois fois plus élevée que dans la souche sauvage. De plus, l'activité de transposition du transposon *Gulliver* est augmentée chez le mutant. Le gène *MUT6* code une protéine similaire aux DEAH-box RNA hélicases qui pourraient participer à la dégradation d'ARN aberrants (ARN double brin) produits par les transgènes et les ET (Okamoto et Hirochika, 2001 ; Wu-Scharf et al., 2001). Les ARN double brin sont coupés par l'enzyme Dicer en petits ARN de 21-22 nucléotides qui sont des indicateurs de l'inhibition post-transcriptionnelle.

Ces différents mécanismes, dont les ET sont les principales cibles, sont plus ou moins interdépendants et sont spécifiques de certains types cellulaires ou de stades de développement (Martinez et Slotkin, 2012). Rigal et Mathieu (2011) ont comparé la complexité du contrôle épigénétique des ET à un "mille-feuille".

## PARTIE II :MATERIELS ET METHODES

### 1. Matériel végétal et conditions de culture

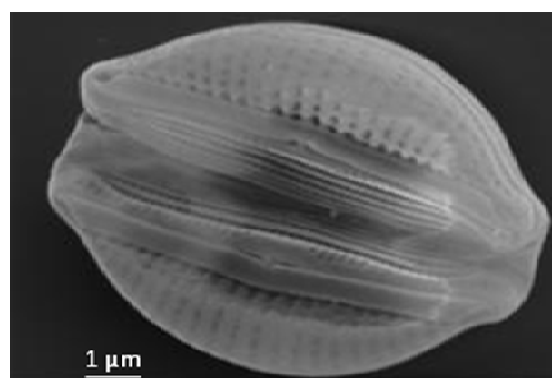
#### 1.1. La diatomée *Amphora acutiuscula*

La diatomée *Amphora acutiuscula* provient de la côte Sud-est du Viêt Nam au climat tropical. Elle a été collectée sur le site de Càn Gio situé dans une mangrove (Fig. 17) et isolée au laboratoire Mer, Molécules, Santé en 2005 par T.L.N. Nguyen-Deroche (2008). Elle est référencée dans la collection de microalgues de l'université de Nantes sous le numéro NCC216. Cette diatomée possède un frustule en silice, rigide et finement ornementé. Du fait de la symétrie bilatérale du frustule, cette espèce appartient au groupe des pennales (ou pennées) (Fig. 18). Ses dimensions sont de 28-32  $\mu\text{m}$  de longueur et 6-8  $\mu\text{m}$  de largeur.

**Figure 17 : Localisation de Càn Gio (latitude: 10°40'09''; longitude: 107°00'59'') sur la côte Sud-est du Viêt Nam.**



**Figure 18 : Photographie de microscopie électronique à balayage de la microalgue marine *Amphora acutiuscula* (cliché de P. Gaudin et T.L.N. Nguyen-Deroche).**



## **1.2. Milieu de culture**

Les cultures axéniques d'*Amphora acutiuscula* sont réalisées dans de l'eau de mer artificielle (EMA) à teneur réduite en calcium (0,25 mM  $\text{CaCl}_2$  au lieu de 9 mM) afin de limiter l'adhésion des cellules entre elles et aux parois de l'Erlenmeyer (Cooskey, 1981; Molino et Wetherbee, 2008). L'EMA est préparée selon le protocole de Harrison et al. (1980) à partir de deux solutions de sels minéraux (Annexe 1) qui sont mélangés après autoclavage (20 min à 110 °C). A l'issue de l'autoclavage, le milieu de culture est stocké en chambre froide à l'abri de la lumière durant quelques jours avant utilisation. La salinité de l'EMA est de 3 ‰ et le pH est de 8,2. Les concentrations en cuivre et en zinc de l'EMA sont de 0,15  $\mu\text{M}$  et 0,25  $\mu\text{M}$ , respectivement (Harrison et al., 1980). L'axénie des cultures est obtenue et entretenue par des traitements antibiotiques et antifongiques réguliers (Annexe 2).

## **1.3. Conditions de culture et de stress**

*A. acutiuscula* est cultivée dans des fioles Erlenmeyer de 500 mL contenant un volume de 250 mL de l'EMA appauvrie en calcium. Ces fioles ne sont ni aérées ni agitées (culture en batch). Les cultures sont maintenues en conditions contrôlées, sous un éclairage de 300  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ , une photopériode de 14 h et une température de  $24 \pm 1$  °C. La mesure du niveau d'éclairage est effectuée avec une sonde quantamétrique 4 $\pi$  (Walz, Allemagne) immergée dans le milieu de culture et reliée à un quantamètre (LiCor, UK).

### ***Stress thermiques et métalliques***

Des cultures axéniques d'*A. acutiuscula* maintenues à 24 °C, ayant chacune une densité initiale de  $10^4$  cellules / mL, ont été soumises à des chocs thermiques au 5<sup>ème</sup> jour de culture, soit en fin de phase exponentielle de croissance (Hermann, 2011). Les cultures ont été placées dans un bain thermostaté à 16 ou 32 °C pendant une courte durée (30 min – 2 h – 5 h – 8 h). Elles n'ont été ni agitées ni aérées pour éviter toute contamination. De même, au 5<sup>ème</sup> jour de culture, pour les stress métalliques, l'EMA a été enrichie en cuivre ou en zinc à partir de solutions mères préparées au laboratoire ( $\text{CuCl}_2$  10 mM et  $\text{ZnCl}_2$  10 mM). Dans ce travail, les concentrations en Cu ajouté (3  $\mu\text{M}$ ) et en Zn ajouté (20  $\mu\text{M}$ ) ont été choisies suite à des essais préliminaires qui avaient pour but de déterminer une gamme de concentrations affectant la croissance d'*Amphora acutiuscula* sans l'inhiber complètement (Nguyen-Deroche, 2008).

## **2. Croissance d'*Amphora acutiuscula***

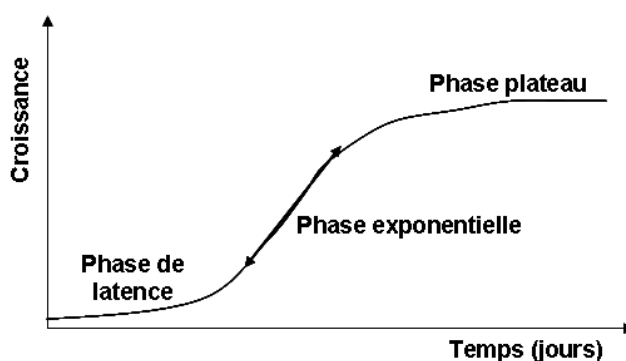
### **2.1. Estimation de la densité cellulaire**

Afin de déterminer la densité cellulaire des cultures de la diatomée *A. acutiuscula*, les cellules sont dénombrées à l'aide de l'hématimètre Neubauer. Avant chaque comptage, les cultures de diatomées ont été homogénéisées pendant 10 min à l'aide d'un agitateur magnétique. Les cellules d'*Amphora* ont tendance à s'agglomérer rendant le dénombrement difficile (Cooksey, 1981). Pour dissocier les cellules les unes des autres (rupture des liaisons hydrogènes au niveau des polysaccharides externes), 1 mL de culture homogène a été mélangé avec 1 mL d'acide HCl 6 M dans un microtube. Ce mélange a été ensuite agité à l'aide d'un micro-barreau aimanté pendant 10 à 15 min. Le calcul de la concentration cellulaire de la culture tient compte de cette dilution.

Toutes les densités cellulaires sont la moyenne de trois comptages différents, dans la zone quadrillée de l'hématimètre de volume 0,1 mm<sup>3</sup>.

### **2.2. Courbe de croissance**

Les courbes de croissance de la diatomée ont été établies en fonction du temps (jours). L'évolution de la densité cellulaire en fonction du temps lors d'une culture en "batch" se traduit par une courbe de forme sigmoïdale (Fig. 19). La croissance se déroule en trois phases :



**Figure 19 : Courbe de croissance théorique des diatomées cultivées en batch en fonction du temps.**

- Phase de latence.
- Phase d'accélération ou phase exponentielle de croissance.
- Phase plateau où la croissance se stabilise après avoir atteint une densité maximale. Cette dernière phase a lieu lorsque la quantité de nutriments disponible diminue. Le suivi des cultures a été réalisé pendant 5 à 6 jours jusqu'à obtention de la phase plateau.

### **2.3. Récolte de la biomasse**

A la fin de chacun des stress appliqués, la biomasse d'*A. acutiuscula* est récoltée par centrifugation à 900 g pendant 15 min et à 4 °C avec une accélération et une décélération lentes pour éviter l'éclatement des cellules. Le culot cellulaire est récupéré dans un tube Falcon de 50 mL. Ce tube est immédiatement immergé dans l'azote liquide pendant 10 min avant conservation à -80 °C jusqu'à l'extraction de l'ADN génomique ou des ARN totaux.

## **3. Activité photosynthétique**

### **3.1. Mesures de l'activité photosynthétique par oxymétrie**

#### ***Extraction et dosage de la chlorophylle a***

Le dosage de la Chl *a* est effectué selon la méthode de Speziale et al. (1984). Un volume de 20 mL de culture a été prélevé pour extraire les pigments chlorophylliens. Ce volume est filtré sous vide partiel sur un filtre GF/C (Whatman). La biomasse des microalgues sur ce filtre peut être conservée à -20 °C. Le filtre est immergé dans 3 mL de diméthyl-formamide (DMF) dans une fiole en verre opaque afin d'éviter la photolyse des pigments. La fiole est ensuite agitée à l'aide d'un agitateur rotatif à température ambiante durant une heure puis 1,5 mL de solution sont centrifugés pendant 5 min à 4 °C et 10 000 g permettant d'éliminer les débris cellulaires. Les densités optiques (DO) du surnageant aux longueurs d'onde de 630 nm, 664 nm et 750 nm sont mesurées. La concentration en Chl *a* est calculée par la formule suivante :

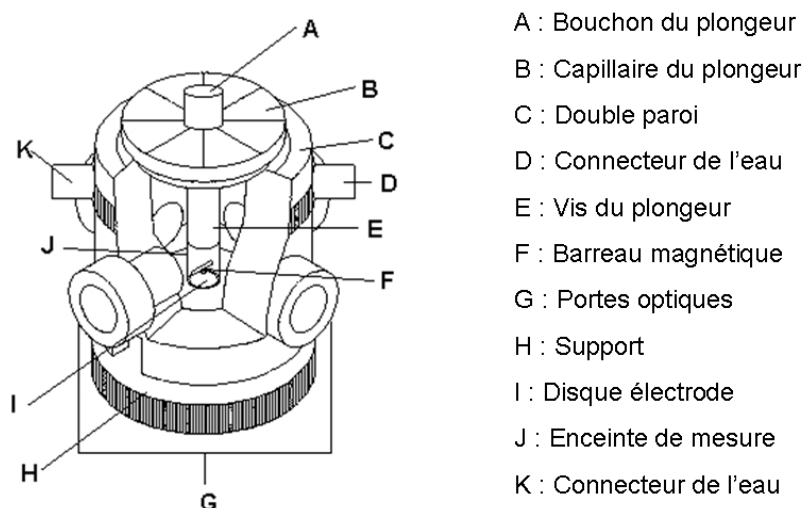
$$\text{Chl}a = \frac{[11,47(\text{DO}_{664} - \text{DO}_{750}) - 0,4(\text{DO}_{630} - \text{DO}_{750})] \cdot V_s}{V_c}$$

où la concentration en Chl *a* est obtenue en  $\mu\text{g mL}^{-1}$ .  $\text{DO}_{630}$ ,  $\text{DO}_{664}$  et  $\text{DO}_{750}$  sont respectivement les densités optiques mesurées aux longueurs d'onde de 630 nm, 664 nm et 750 nm.  $V_s$  est le volume de DMF utilisé pour l'extraction (ici 3 mL).  $V_c$  est volume de culture de microalgues utilisé (ici 20 mL).

#### ***Mesure de l'intensité photosynthétique***

Dans cette étude, l'intensité photosynthétique (IP) a été mesurée par oxymétrie qui est une technique facile à mettre en œuvre et qui est couramment utilisée dans notre laboratoire.

L'émission de dioxygène ( $O_2$ ) a été mesurée à l'aide d'un oxymètre composé d'un boîtier (Hansatech LTD, UK) de contrôle couplé à une électrode de Clark (Delieu et Walker, 1972) ou électrode à oxygène. L'électrode à oxygène est composée d'une anode en argent et d'une cathode en platine, recouvertes d'une membrane perméable à l'oxygène. Un volume de 1,5 mL de suspension de diatomées prélevé dans la culture est placé dans la cellule de mesure (Fig. 20). Cette enceinte de mesure est partiellement fermée dans sa partie supérieure par un bouchon traversé par un capillaire réduisant le contact avec  $O_2$  atmosphérique. La suspension est agitée par un micro-barreau aimanté pour maintenir la teneur en  $O_2$  homogène dans la cellule de mesure. Les diatomées sont maintenues à température constante (température de culture ou de stress) par circulation d'eau thermostatée dans la double paroi de la cellule de mesure. L'électrode à oxygène est connectée à un micro-galvanomètre qui affiche l'intensité du courant électrique en mV, proportionnelle à la concentration d' $O_2$  présent dans la cellule de mesure (1 mg  $O_2$   $L^{-1}$  correspond à 100 mV).



**Figure 20 : Vue éclatée de la cellule de mesure de l'oxymètre.**

Avant chaque série de mesures, l'oxymètre est calibré avec de l'azote gazeux (zéro) et de l'air ambiant (teneur maximal en  $O_2$ ), à la température désirée (16, 24 ou 32 °C). L'évolution de la teneur en  $O_2$  dans la cellule de mesure est enregistrée en fonction du temps. La pente de la courbe qui en résulte permet de déterminer l'activité photosynthétique AP (exprimée en mg  $O_2$   $L^{-1}$   $min^{-1}$ ). L'intensité photosynthétique (IP) rapportée à un mg de Chl *a* est ensuite calculée à l'aide de la relation ci-dessous :

$$IP = \frac{AP \cdot 60 \cdot 10^3}{32 \cdot [Chl a]}$$

IP : intensité photosynthétique en  $\mu\text{mol O}_2 \text{ h}^{-1} \text{ mg}^{-1} \text{ Chl } a$ .

AP : activité photosynthétique en  $\text{mg O}_2 \text{ L}^{-1} \text{ min}^{-1}$ .

[Chl *a*] : teneur en Chl *a* ( $\text{mg L}^{-1}$ ).

32 est la masse molaire de l'oxygène ; 60 : conversion en heure ;  $10^3$  : conversion des mg en  $\mu\text{g d'O}_2$ .

### **3.2. Mesure de la fluorescence de la chlorophylle *a***

#### ***Principe du fluorimètre modulé***

Afin d'étudier la réponse de l'appareil photosynthétique dans les différentes conditions expérimentales, les paramètres de la fluorescence chlorophyllienne ont été mesurés par fluorimétrie modulée. Cette technique est non intrusive et fournit des informations rapides, fiables et reproductibles sur l'appareil photosynthétique et plus particulièrement sur le photosystème II (PSII). Les paramètres de fluorescence sont mesurés *in vivo* directement sur la suspension de microalgues et sont le reflet inversé des événements photochimiques qui ont lieu au niveau du PSII. Le principe de la fluorimétrie modulée repose sur l'utilisation de deux sources de lumière permettant de détecter des variations de la fluorescence.

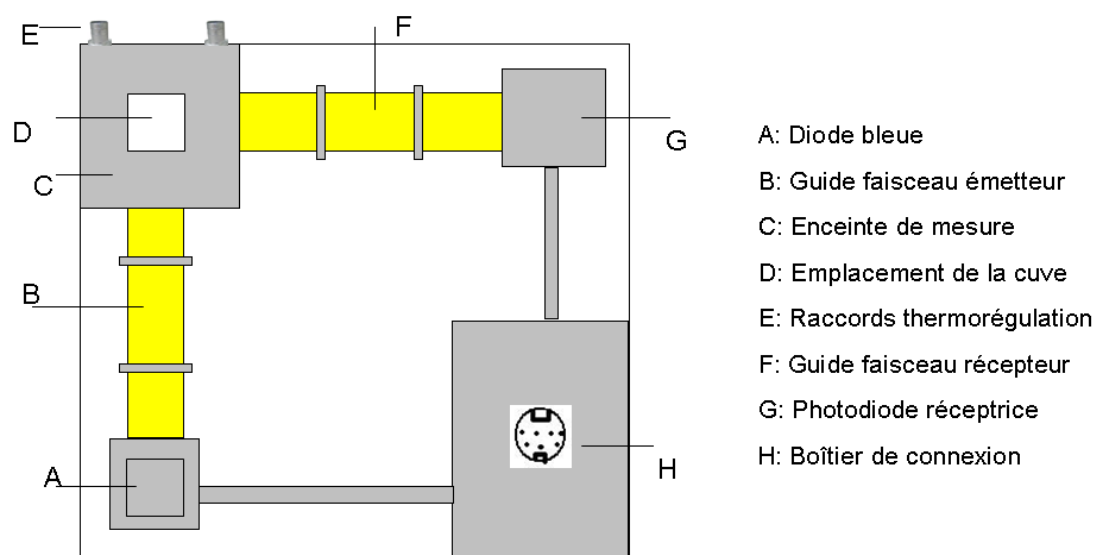
- La première source lumineuse de faible puissance (n'induisant pas la photosynthèse) émise sous forme d'un signal modulé avec une certaine fréquence, est utilisée pour mesurer la fluorescence. Cette lumière modulée ou lumière analytique est produite par des diodes électroluminescentes bleues (LED). Elle est couplée à un récepteur permettant de mesurer et d'enregistrer les variations du signal de fluorescence émis par la Chl *a*. Dans ces conditions, la lumière actinique induisant la photosynthèse n'influence pas la détection de la fluorescence.
- La seconde source est actinique. Son intensité est variable avec plusieurs niveaux d'éclairement prédéfinis. Cette source lumineuse est utilisée pour induire la photosynthèse et / ou changer l'état du PSII.



### ***Fluorimètre modulé FMS-1***

Le fluorimètre FMS-1 (Hansatech) a été modifié par Rech et al. (2003) et utilisé avec succès pour effectuer des mesures avec des suspensions de microalgues marines peu concentrées (Fig. 21). Il est couplé à un ordinateur permettant d'enregistrer les valeurs et de calculer les rapports correspondant aux paramètres de fluorescence. La lumière modulée est émise par une diode bleue au gallium (Nichia NSPB-500S) pour induire la fluorescence qui est ensuite captée par une photodiode réceptrice. La lumière actinique est conduite par une fibre optique à l'échantillon placé dans une cuve en verre bimiroitée située au croisement des guides optiques des lumières émettrice et réceptrice. La lumière actinique active la photosynthèse, l'émission d'un flash de lumière saturante permet de fermer tous les centres réactionnels du PSII et de mesurer la fluorescence maximale.

L'échantillon est agité par un barreau aimanté placé dans la cuve pour assurer l'homogénéité de la suspension. L'enceinte de mesure est thermo-régulée (température adaptée à chaque échantillon d'algues :  $16$  et  $32 \pm 1$  °C pour les stress thermiques ou  $24 \pm 1$  °C pour le témoin). Le logiciel qui pilote le fluorimètre affiche sur l'écran de l'ordinateur les paramètres de fluorescence et une zone graphique délivre en temps réel les différentes valeurs ou paramètres enregistrés.



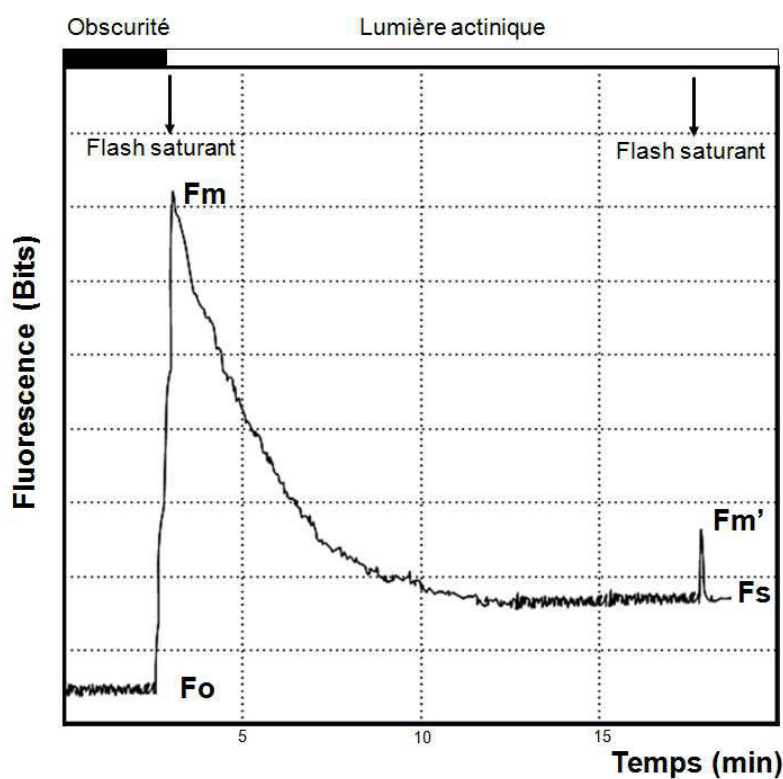
**Figure 21 : Vue de dessus de l'extension du fluorimètre modifié par Rech et al. (2003).**

La lumière émise par la diode émettrice est conduite à l'enceinte de mesure par le guide faisceau émetteur. Le signal de fluorescence est capté par la photodiode (photodétecteur) placée à l'extrémité du guide faisceau récepteur. Les signaux électriques sont envoyés et reçus via le boîtier de connexion (Rech *et al.*, 2003).

## Mesures

Les suspensions d'algues sont placées dans une cuve bi-miroitée, thermo-régulée et agitée au croisement des sources lumineuses.

Avant chaque mesure, les diatomées sont placées 10 min à l'obscurité. Pendant cette période d'adaptation, tous les centres réactionnels du PS II sont ouverts. L'émission de lumière modulée va permettre de mesurer la fluorescence minimale  $F_o$  (Fig. 22). Un flash saturant de lumière blanche est alors envoyé conduisant à la fermeture des centres réactionnels du PSII qui deviennent alors inactifs. La fluorescence maximale  $F_m$  peut alors être enregistrée. A partir de ces deux paramètres, la fluorescence variable,  $F_v$  qui correspond à la fraction de l'énergie photonique absorbée convertible en énergie chimique, est calculée selon l'équation :  $F_v = F_m - F_o$ . Le calcul du rapport  $F_v/F_m$  correspond alors à une mesure du rendement quantique maximum des réactions photochimiques au niveau du PSII.



**Figure 22 : Évolution de la fluorescence chlorophyllienne d'une suspension d'*A. acutiuscula* à 24 °C.**

La fluorescence minimale  $F_o$  est mesurée à l'obscurité. La fluorescence maximale  $F_m$  est ensuite obtenue après émission d'un flash de lumière saturante. A la lumière, la fluorescence stabilisée  $F_s$  est mesurée et la fluorescence maximale  $F_m'$  est obtenue à la suite d'un flash saturant. Le système couplé à un ordinateur permet de calculer en temps réel le rapport  $F_v/F_m$  et  $\Phi_{PSII}$ .

L'échantillon est ensuite éclairé pendant 10 min avec de la lumière actinique d'intensité  $300 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , la fluorescence de base ( $F_s$ ) peut être mesurée à la fin de cette période d'éclairement. L'émission d'un flash saturant, identique à celui utilisé pour la mesure de  $F_m$  permet d'obtenir  $F_m'$  (Fig. 22), la fluorescence maximale à la lumière, après 10 min d'adaptation à la lumière actinique. Il est alors possible de calculer le paramètre  $\Phi_{\text{PSII}}$ , ou rendement quantique effectif du PSII d'après l'équation :  $\Phi_{\text{PSII}} = (F_m' - F_s) / F_m'$ .

#### **4. Analyses statistiques**

Les valeurs moyennes de trois réplicats réalisés dans les mêmes conditions sont présentées avec les écarts types dans les tableaux 15, 16 et 17. D'abord, le logiciel Excel a fourni un moyen d'analyse statistique sur la moyenne et l'écart à la moyenne. Puis toutes les données ont été transférées et traitées dans le logiciel statistique R pour effectuer d'autres analyses. Les comparaisons entre différents groupes d'échantillons sont réalisées par ANOVA (two way analysis of variance). La différence entre les groupes est statistiquement significative quand les différences des valeurs moyennes entre les groupes expérimentaux sont plus grandes que les différences obtenues au hasard. Le niveau de confiance est donné par la formule  $100 \times (1 - \alpha)$  ou 95 % ; où  $\alpha$  est égal 0,05, c'est-à-dire que un vingtième de probabilité d'erreurs est acceptable. Autrement dit, le coefficient de variation ( $p$ ) ne doit pas dépasser 5 %. Une analyse de variance ne permet pas d'identifier quelle moyenne particulière est différente des autres. On a besoin d'une approche plus formelle pour considérer les échantillons par paires et trouver des moyennes qui sont significativement différentes. Le test de HSD Tukey (HSD pour *Honestly Significant Difference*) a été choisi pour effectuer des comparaisons multiples. Il effectue un test  $t$  séparé pour chaque paire d'échantillons à comparer en faisant les ajustements de la  $p$ -value.

## **5. Techniques de biologie moléculaire**

### **5.1. Extraction des acides nucléiques**

#### ***Extraction d'ADN génomique***

L'ADN génomique d'*A. acutiuscula* a été extrait selon le protocole modifié de Doyle et Doyle (1990) à partir de 250 mL de cultures axéniques en phase exponentielle de croissance. Les cellules ont été collectées par centrifugation à 900 g avec accélération et décélération lente à 4 °C pendant 10 min puis congelées à -80 °C. Le culot cellulaire est broyé dans un mortier en présence d'azote liquide. La poudre est récupérée dans un tube contenant 5 mL de tampon d'extraction ( $\beta$ - mercaptoéthanol 0,2 %; CTAB 2 %; EDTA 20 mM, pH 8; NaCl 1,4 M et Tris 100 mM, pH 8) préalablement chauffé à 60 °C. L'ensemble est incubé pendant 30 min à 60 °C en retournant doucement le tube de temps en temps. Après avoir ajouté 5 mL de chloroforme, le mélange est homogénéisé doucement par retournement durant 5 min puis centrifugé pendant 15 min à 8000 g à 4 °C. La phase aqueuse supérieure est transférée dans un nouveau tube et un volume égal d'isopropanol est rajouté afin de précipiter les acides nucléiques. Après une incubation de 2 h à -20 °C, le mélange est centrifugé pendant 15 min à 8000 g à 4 °C. Le culot d'ADN est récupéré avec un cône de pipette 1000 et subit un lavage dans un tube d'Eppendorf contenant 500  $\mu$ L d'éthanol 70 %. Après centrifugation de 8000 g à 4 °C pendant 15 min et élimination du surnageant, le culot d'ADN est séché sur paillasse à l'air libre et repris dans 80  $\mu$ L de tampon TE (EDTA 1 mM, pH 8 et Tris-HCl 10 mM, pH 8). Un traitement final a été réalisé par ajout de 1,5  $\mu$ L de solution de RNase (4 mg/mL). Le tout est mélangé doucement puis placé à 37 °C pendant une heure. Les ADN génomiques sont conservés à -20 °C.

#### ***Extraction des ARN totaux***

L'ensemble du matériel utilisé est traité à l'éthanol 70 % et immergé dans du peroxyde d'hydrogène H<sub>2</sub>O<sub>2</sub> à 3 % durant 15 min, suivi de deux rinçages à l'eau DEPC 0,1% (diéthyl pyrocarbonate) afin d'éliminer un maximum de ribonucléase (RNase). Le reste du matériel (mortier, pilon, spatule) est stérilisé au four Pasteur pendant 3 heures à 180 °C. Deux protocoles d'extraction ont été utilisés, le premier au TriReagent pour toutes les extractions dites de routine et pour les échantillons servant aux transcriptions inverses pour vérifier la présence ou l'absence des transcrits, et la deuxième à l'aide d'un kit (Sigma) pour tous les échantillons de PCR en temps réel.

Les culots cellulaires d'*A. acutiuscula* sont broyés dans un mortier en présence d'azote liquide pour éviter une réactivation des enzymes dégradant les acides nucléiques et principalement les ARN. Le broyat est divisé à raison d'environ 50 mg par tubes et conservé dans l'azote liquide puis stockés à -80 °C pour de futures extractions. 1 mL de Tri Reagent ainsi que 50 µL de bromoanisole (1-bromo-4-méthoxybenzène) sont ajoutés au 50 mg de broyat. La solution est homogénéisée au vortex pendant 15 s et placée sur glace pendant 15 min. Cette étape permet une dissolution des cellules et la libération des acides nucléiques ainsi qu'une dénaturation des ADN à pH acide. Puis une centrifugation de 15 min à 4 °C et 12 000 g permet de séparer les débris cellulaires des ARN restant en phase aqueuse. Afin de précipiter les ARN, de l'isopropanol est ajouté à volume égal à la phase aqueuse et l'ensemble est incubé sur glace durant 10 min et une centrifugation à 4 °C et 12 000 g durant 5 min permet d'obtenir un culot d'ARN. Toute trace d'isopropanol est éliminée par un lavage à l'éthanol 75 %. Ce dernier est éliminé à la suite d'une centrifugation de 5 min à 4 °C et 6 000 g. Le culot d'ARN est séché à l'air libre et enfin mis en suspension dans 20 µL d'eau DEPC. Les ARN totaux sont finalement conservés à -80 °C pour toute analyse ultérieure.

La deuxième méthode d'extraction se fait grâce à un kit (Spectrum, Plant Total RNA Kit, Sigma). Les microalgues filtrées sont broyées dans un mortier froid avec 1 mL de Lysis Solution (pour 100 mg de microalgues). Le tout est incubé 5 min à 56 °C dans un microtube puis centrifugé 3 min à 15 000 g à 21 °C (toutes les autres centrifugations se feront dans les mêmes conditions) afin d'éliminer les débris cellulaires. Les débris résiduels sont filtrés sur une colonne spécifique du kit. Au surnageant sont ajoutés 750 µL de Binding Solution et le tout est passé à travers la colonne d'accrochage des ARN. Un traitement DNase (10 µL de DNase + 70 µL de digestion Buffer, Sigma) est réalisé directement sur colonne 15 min à température ambiante. Après lavage, la colonne est placée dans un nouveau tube collecteur et 40 µL d'Elution Solution sont ajoutés afin de récupérer les ARN en solution.

## **5.2. Contrôle des extractions**

La quantité et la qualité de l'ADN génomique et des ARN sont mesurées par densité optique à 260 nm (une unité de DO correspondant à 50 µg/mL d'ADN double brin et 40 µg/mL pour l'ARN) au moyen d'un spectrophotomètre. La pureté des produits d'extraction est évaluée grâce au rapport  $DO_{260} / DO_{280}$  entre 1,8 et 2 pour l'ADN et

autour de 1,6 pour les ARN. Pour cela, 1 µL d'extrait d'acides nucléiques est dilué à 1 % avec de l'eau ultra pure puis l'absorbance est mesurée à 260 nm (absorbance des acides nucléiques) et 280 nm (contamination protéique).

La présence et la taille des acides nucléiques sont analysées sous UV après électrophorèse sur gel d'agarose à 1 % en présence de bromure d'éthidium. Les ARN ribosomiaux sont majoritaires (~ 80 %) dans les ARN totaux et sont visualisés sous la forme de deux bandes principales (28S et 18S). Une "bande" supplémentaire de petite taille indique une dégradation des ARN totaux.

### **5.3. Amplification par réaction de polymérisation en chaîne (PCR)**

Pour chaque PCR, un volume final de 25 µL ou 50 µL de milieu réactionnel a été utilisé comportant du tampon Taq 1X (Promega, Madison, WI), du MgCl<sub>2</sub> 2,5 mM, des dNTP 0,2 mM, de la *Taq* polymérase à raison de 0,5 U pour 25 µL (Promega, Madison, WI), des amorces 1 µM. Pour les PCR sur ADN génomique, 100 à 150 ng d'ADN (pour 25 µL) ont été utilisés comme matrice. Respectivement 200 ng et 400 ng d'ADNc ont été utilisés comme matrice pour rechercher les transcrits. Le programme PCR a suivi le modèle suivant, avec des variations de durée d'élongation ou de température d'hybridation : 94°C x 3 min, [94 °C 1 min ; T<sub>m</sub> hybridation °C 1 min ; 72 °C 30 s] x 40 cycles, 72 °C x 10 min. L'amplification PCR a été utilisée à des fins diverses :

- détecter la présence de différents gènes : ménage, marqueurs de stress et éléments transposables dans le génome de la diatomée *A. acutiuscula*.
- mettre en évidence la présence de transcrits du gène MLE chez *Amphora acutiuscula*, ainsi que des gènes de réponse aux stress. Cette technique a été réalisée avec de l'ADN complémentaire.
- vérifier l'insertion du gène d'intérêt dans les plasmides des bactéries compétentes au cours du clonage. Des PCR sur les ADN plasmidiques ont été réalisées avec les amorces situées sur le fragment d'intérêt avec les amorces ayant permis son amplification. Les programmes d'amplification varient de 35 à 40 cycles avec une dénaturation initiale de 5 min à 94 °C. La température d'hybridation est variable selon les amorces utilisées et sont répertoriées (Tab. 1).

**Tableau 1 : Tableau récapitulatif des couples d'amorces dégénérées ou spécifiques des différents gènes étudiés.**

Les amorces *MLE5A* et *MLE3A* pour détecter la présence des MLE dans le génome d'*A. acutiuscula*. Les amorces pour mettre en évidence la présence des gènes de ménage (*Actine* et *18s*) et de leurs transcrits. Les amorces pour mettre en évidence la présence des gènes de réponse au stress (*MnSOD* et *Hsp70*) et leur transcrits. Tm : température d'hybridation des amorces. H = A/T/C, R = G/A, Y = T/C, D = G/A/T.

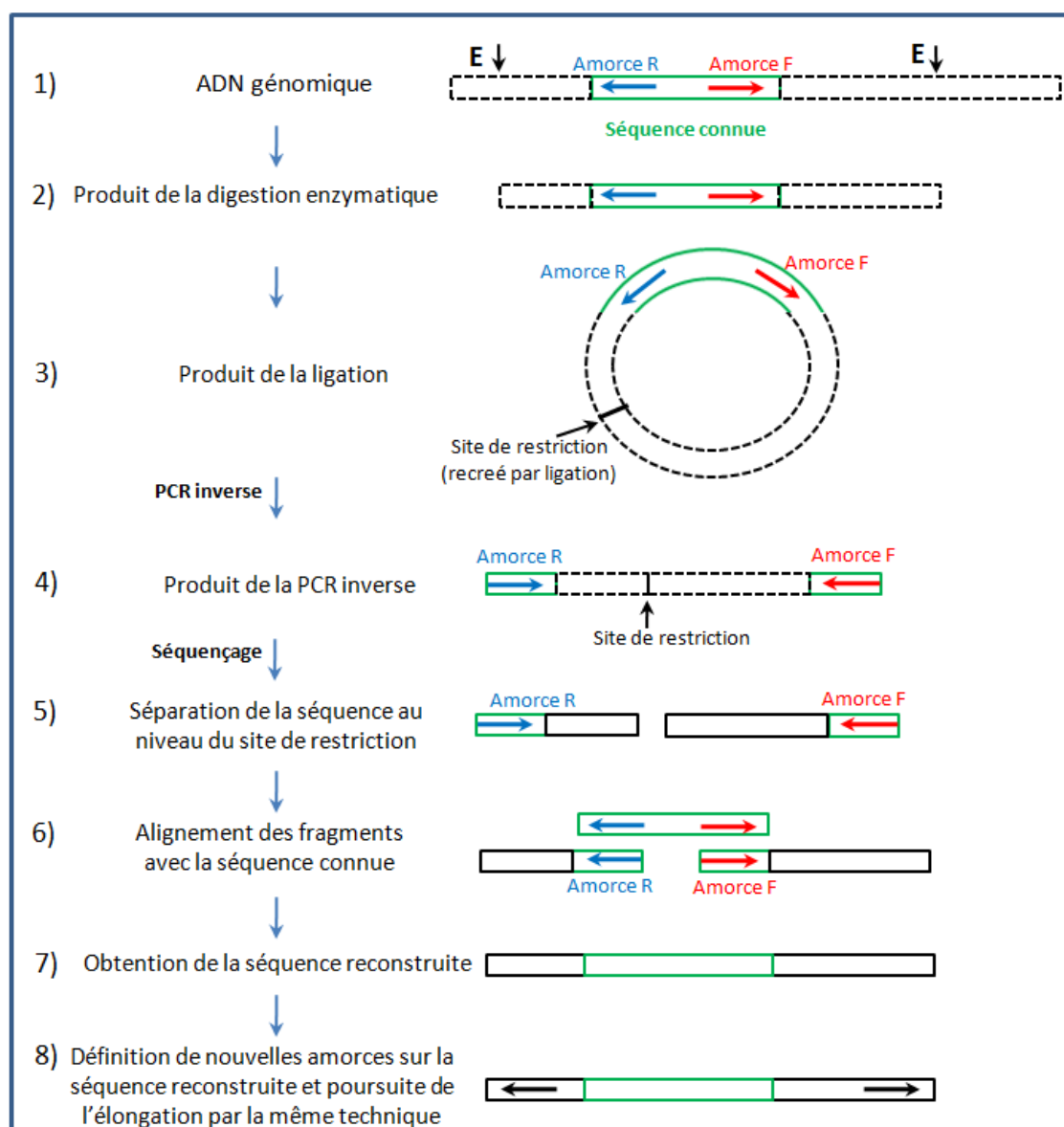
Gène	Amorces	Séquence (5' → 3')	Fragment amplifié	Tm (°C)
<i>MLE</i>	<i>MLE5AM</i>	ATH GAT GAR AAR TGG TTC GCA TTR	380 pb	47
	<i>LE3A</i>	TCY TGY TGD AT		
	<i>AcutFcAc</i>	TAC ATT GGC AAA GTG ACG TTC TTGTCC	140 pb	59
	<i>utRe</i>	TCT CGG TCT GTG TTT CGA T		
	<i>AcutF9ext</i> <i>AcutR8ext</i>	ATATCGATGCAATGGATGAACGTTGATTCTGCTCTCATGTTTGAG	720 pb	55
<i>Actine</i>	<i>AaactinF</i>	ATATCGATGCAATGGATGAACGTC AATAC	900 pb	55
	<i>AaactinR</i>	ATCTGATTGAGCTT		
<i>18s</i>	<i>RibFRib</i>	CCTCATGCCGTCATTCGTCTTGACCTGAC	150 pb	62
	<i>R</i>	GGCAACGAAGCAGAGGGATTCC		
<i>MnSOD</i>	<i>F-SODspR-SODsp</i>	GGG GGA GTA TGG TCG CAA GGCTCA GTG TAG CGC GCG TGC GGC	370 pb	61
		ATC TGT CGC CTA ACT CGTAA CTC GTC GGT TCC TTC TG		
<i>Hsp70</i>	<i>Fhsp70R</i>	CAACGWCAGGCWACCAAGCCATTGGTM	200 pb	52
	<i>hsp70</i>	GCCTTKACTTC		
	<i>NFhsp70</i> <i>NRhsp70</i>	GGCWACCAAGGATGCTGGRAAGCCTTKA CTTCAAAYACTCC	190 pb	53

#### **5.4. PCR inverse et reconstruction de séquences**

La PCR inverse (iPCR) permet de connaître les régions flanquantes inconnues d'une séquence connue d'ADN. L'ADN génomique est digéré par une ou plusieurs enzymes de restriction, puis les fragments sont circularisés et servent de matrice pour une réaction de PCR avec des couples d'amorces "dites "sortantes" situées dans la région de la séquence connue (Fig. 23). Les amplifications obtenues sont clonées et séquencées, permettant de reconnaître les séquences des régions flanquant l'insert sur sa droite et sa gauche. 500 ng d'ADNg sont digérés en présence de 5 à 10 U d'enzymes de restriction (Tab. 2). La circularisation des produits de digestion est réalisée en présence de 30 µL de tampon de réaction et 100 U de T4 DNA Ligase dans un volume final de 150 µL. Le mélange est incubé pendant 5 h à 37 °C puis les enzymes sont inactivées à 65 °C pendant 10 min. 2 µL de produits circularisés sont utilisés comme matrice pour l'inverse PCR (dans les mêmes conditions que celles décrites précédemment) avec différentes amorces (Tab. 2 et Fig. 23).

En parallèle à l'allongement, il convient de vérifier, par PCR "classique", dans les mêmes conditions que citées plus haut, les séquences allongées obtenues. Pour cela, des amorces dites "entrantes" sont désignées sur les séquences reconstruites (Fig. 24 et Tab. 3). L'objectif étant ici d'obtenir, sur ADNg, une amplification pleine taille correspondant à ce qui a été obtenu par iPCR et mettre en évidence l'ORF, les UTR et ITR. Ces amplifications PCR sont ensuite alignées et comparées avec les séquences reconstruites





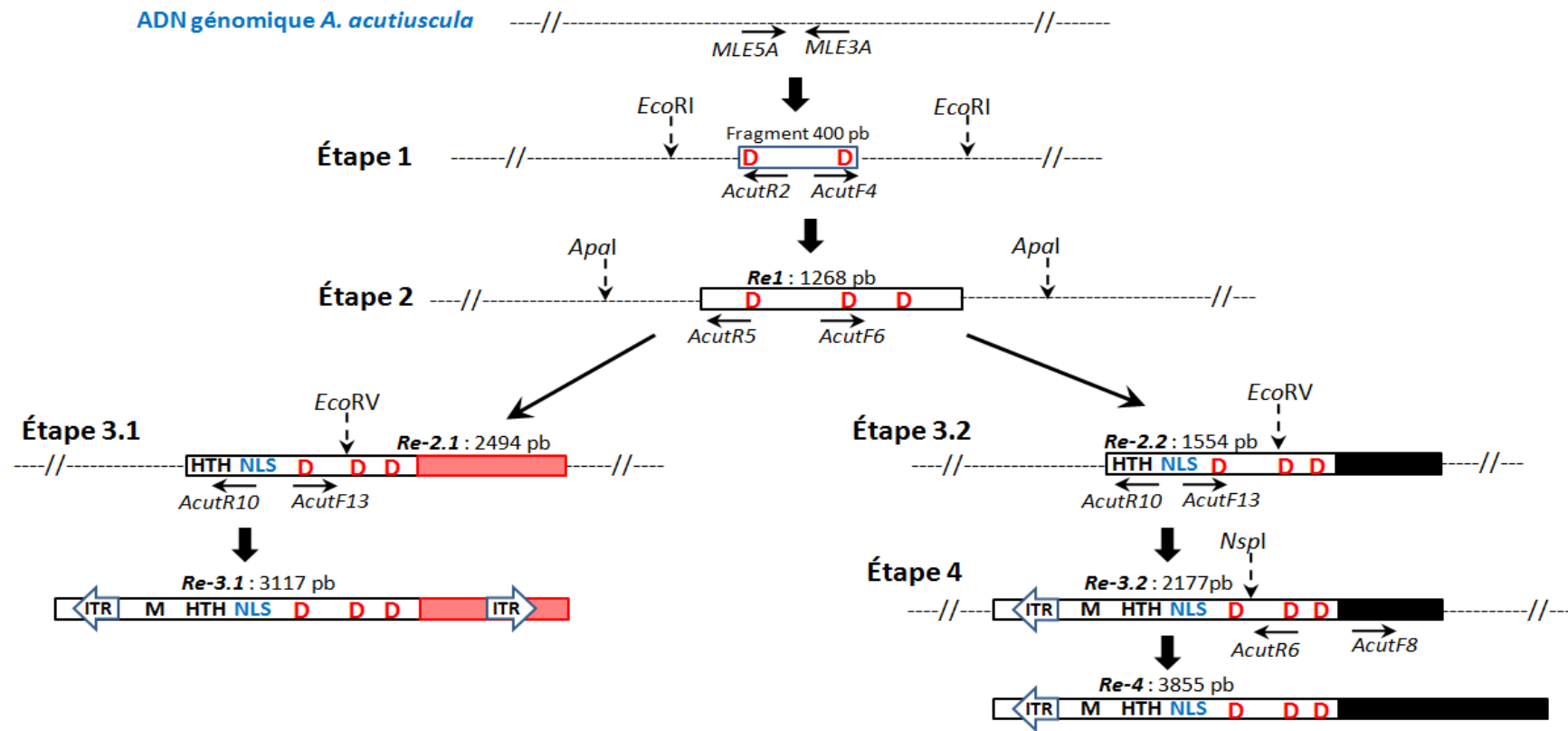
**Figure 23 : Allongement d'une séquence connue par PCR inverse.**

1) et 2) : l'ADN génomique de la diatomée *A. acutiuscula* est digéré à l'aide d'une enzyme de restriction qui coupe, ici, à l'extérieur de la séquence connue. 3) et 4) : Les produits de digestion sont ligués sur eux-mêmes et servent de matrice pour la PCR inverse. 5) : après séquençage, l'analyse des séquences permet de localiser les amorces et le site de restriction. C'est au niveau de ce site que la séquence est séparée en deux fragments. 6) et 7) : Chaque fragment est alors aligné avec la séquence connue : la partie commune entre les fragments et la séquence connue permet de vérifier que c'est bien cette dernière qui a servi de matrice à l'allongement de la séquence. La partie située au delà de la séquence connue correspond à l'allongement de la séquence. Une fois les trois fragments assemblés, nous obtenons une séquence reconstruite. 8) la séquence reconstruite est utilisée comme matrice pour définir de nouvelles amorces qui serviront à poursuivre l'allongement de la séquence en utilisant la même technique. E↓: Localisation du site de restriction enzymatique. Les séquences connues sont indiquées en traits pleins. Les séquences inconnues sont indiquées en pointillés.

**Tableau 2 : Tableau récapitulatif des enzymes et amorces utilisées en iPCR.**

Les trois enzymes de restriction (*EcoRI*, *ApaI*, *EcoRV* et *NspI*) et leur site de coupure sont présentés. Nous avons récupéré les amorces pour chaque réaction iPCR ainsi que le nom et taille du produit PCR séquencé pour la reconstruction des fragments finaux *Re-3.1* et *Re-4.2* à partir de la séquence consensus partielle de type MLE (400 pb).

Enzyme et site de coupeure	Nombre de coupeure	Amorces pour PCR inverse	Tm	Nom et taille du produit PCR séquencé
<i>EcoRI</i> 5'...G▼AATT▲C...3'	0	<i>AcutR2ext</i> : GCACACAAGAAC (A/G) TCAC(T/C) TTGC	PCR <sub>1</sub> : 55°C	<i>EcoRIR2F4-1</i> : 1217 pb <i>EcoRIR2F4-2</i> : 1217 pb
		<i>AcutF4ext</i> : TTGG(C/T) AT(T/A) TGGCCATTTGGA(C/T) ATG	PCR <sub>2</sub> : 55°C	
		<i>AcutR2int</i> : CAAGAAC (A/G) TCAC(T/C) TTGCCAATGTA		
		<i>AcutF4int</i> : TGGCCATTTGGATATGTTGG		
<i>ApaI</i> 5'...G▼GGCC▲C...3'	0	<i>AcutR5ext</i> : AAGTACTTCGGACTTCCATTCC	PCR <sub>1</sub> : 55°C	<i>ApaIR5F6-1</i> : 1982 pb <i>ApaIR5F6-2</i> : 1983 pb
		<i>AcutF6ext</i> : TCAATCACCAGACACAAACATC	PCR <sub>2</sub> : 57°C	
		<i>AcutR5int</i> : ACTTCCATTTCCTTACCATATTTGGG		
		<i>AcutF6int</i> : ACACAAACATCAATGATTTGGCATTC		
<i>EcoRV</i> 5'...GAT▲ATC...3'	1	<i>AcutR10ext</i> : TCTCCTACACAGCGCGGTT	PCR <sub>1</sub> : 53°C	<i>EcoRVR10F13-1</i> : 727 pb <i>EcoRIR10F13-2</i> : 728 pb
		<i>AcutF13ext</i> : GTGGATGAAGATGAGATCAAATC	PCR <sub>2</sub> : 55°C	
		<i>Acut10Rint</i> : AGCGCGGTTGTGCCTCAATT		
		<i>AcutF13int</i> : GATCAAATCAAAGAAACATGAGAGAG		
<i>NspI</i> 5'...R▼CATG▲Y...3'	1	<i>AcutR6ext</i> : ATGGCATCTAACCATTCTTCATC	PCR <sub>i</sub> : 55°C	<i>NspIR6F8-1</i> : 2886 pb <i>NspIR6F8-2</i> : 2886 pb
		<i>AcutF8ext</i> : TACATCAAACACAGCTAACTGC	PCR <sub>n</sub> : 57°C	
		<i>AcutR6int</i> : TTCTTCATCATTGTCAGGAATATGGG		
		<i>AcutF8int</i> : TAACTGCGACTGGTCCATCTGGGT		



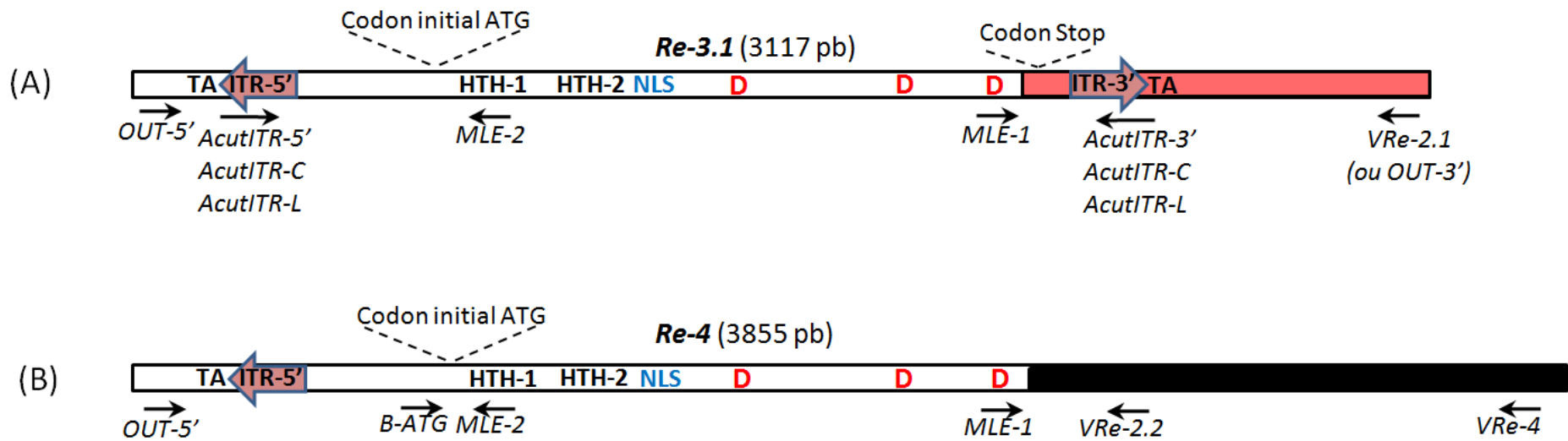
**Figure 24 : Schéma récapitulatif des étapes de la reconstruction des séquences MLE chez *A. acutiuscula*.**

Le fragment partiel de MLE (400 pb) a été détecté par l'utilisation des amorces dégénérées *MLE5A* et *MLE3A*. **Étape 1** : l'ADNg d'*A. acutiuscula* est digéré par l'enzyme *EcoRI* qui ne coupe pas le fragment connu de 400 pb. Les amorces *AcutR2/AcutF4* sont utilisées pour construire un fragment *Re-1* de 1268 pb qui contient également le 3ème D de la triade catalytique. *Re-1* est le point de départ pour l'iPCR de l'étape 2. **Étape 2** : l'ADNg d'*A. acutiuscula* est digéré par l'enzyme *ApaI* qui ne coupe pas le fragment reconstruit *Re-1*. Les amorces *AcutR5/AcutF6* sont utilisées pour reconstruire deux fragments *Re-2.1* (2494 pb) et *Re-2.2* (1554 bp). Ces fragments partagent une partie identique en région 5' contenant les motifs caractéristiques des MLE et une partie différente en région 3'. **Étape 3** : l'ADNg d'*A. acutiuscula* est digéré par l'enzyme *EcoRV* qui coupe à l'intérieur des *Re-2*. Les amorces *AcutR10/AcutF13* sont utilisées pour allonger leur extrémité 5' et permettre de reconstruire les fragments *Re-3.1* et *Re-3.2*. Nous avons réussi à identifier un transposon MLE complet (*Aacutmar1-re*) dans la séquence de *Re-3.1* avec ses ITR à chaque extrémité. **Étape 4** : l'élongation d'extrémité 3' du *Re-3.2* par l'enzyme *NspI* et les amorces *AcutR6/AcutF8*, cela a permis d'obtenir le *Re-4* (3855 pb).

**Tableau 3 : Tableau récapitulatif des couples d'amorces utilisés pour vérifier des fragments reconstruits et amplifier les séquences contenant ITR, les MLE complets.**

Groupe 1 pour vérifier la présence réelle des fragments reconstruits dans le génome *A. acutiuscula*. Groupe 2 et le couple *MLE-1/OUT-3'* dans le groupe 1 servent à amplifier les fragments qui contiennent les ITR-5' et 3', respectivement. Groupe 3 composés des amorces spécifiques et dégénérées pour amplifier les MLE complets dans le génome *A. acutiuscula*. Y = C/T, R = A/G, W = A/T. La position des amorces est déterminée à partir de la première base de chaque séquence reconstruite finale *Re-3.1* et *Re-4* (Fig. 25).

Groupe	Couple d'amorces	Séquence (5' → 3')	Position	Tm	Nom et taille du fragment attendu
1	<i>MLE-1</i> <i>VRe-2.1 (ou OUT-3')</i>	ACA CAA ACA TCA ATG ATA TGG CAT TC AGAGAC CTG TCA TAT TGA AGG	1537-1563 2591-2611	51	<i>MLE1/VRe2.1</i> : 1085 pb
	<i>MLE-1</i> <i>VRe-2.2</i>	ACA CAA ACA TCA ATG ATA TGG CAT TC TTC GAA TGG AAT GAA AGA GTG TGG	1537-1563 2073-2096	55	<i>MLE1/VRe2.2</i> : 559 pb
	<i>B-ATG</i> <i>VRe-4</i>	TTG AAC AGA AGG CAA TGA AAA G ATCGAGACTTGAAAGCCTTATCG	419-440 3741-3753	56	<i>BATG/VRe4</i> : 3303 pb
	<i>OUT-5'</i> <i>MLE-2</i>	AAA TAC GGT GTC GGT GGC AGC CAT CTG TAA AAC TAC CAT GGG GAA GA	58-78 624-649	56	<i>OUT5'/MLE2</i> : 592 pb
3	<i>AcutITR-5'</i> <i>AcutITR-3'</i>	TAC TCC CTC TGG TCC AAA ATT A TAC CCC CTC TGG TCC AGA ATT A	274-295 2383-2403	58	<i>Aacutmar1</i> : 2130 pb ; <i>Aacutmar2</i> : 1700 pb et <i>Aacutmar3</i> : 1000 pb
	<i>AcutITR-C</i>	TAC YCC CTC TGG TCC ARA ATT A	274-295	58	
	<i>AcutITR-L</i>	TAC YCC CTC TGG TCC ARA ATT AWW CAT CA	274-302	64	



**Figure 25 : Schéma de l'emplacement des amorces permettant de vérifier la présence réelle des séquences reconstruites dans le génome d'*A. acutiuscula*.**

(A) Fragment reconstruit *Re-3.1* qui contient un transposon MLE complet bordé par les deux ITR. (B) Dans le fragment reconstruit *Re-4*, on n'identifie pas l'ITR-3'. Les motifs HTH (Hélice-Tour-Hélice) en noir et NLS (Signal de Localisation Nucléaire) en bleu sont trouvés dans le domaine de liaison à l'ADN. Le domaine catalytique est caractérisé par la triade DDD en rouge.

### **5.5. RT- PCR**

Avant de réaliser une PCR après transcription inverse (Reverse transcriptase-PCR : RT-PCR), les ARN totaux ont subi un traitement à la DNase pour éliminer toute trace d'ADN génomique contaminant. 1,5 - 2 µg d'ARN sont ajoutés à un mélange réactionnel composé de 1 µL de tampon 10 X RQ1 (400 mM Tris-HCl pH 8, 100 mM MgSO<sub>4</sub>, 10 mM CaCl<sub>2</sub>), et 2 U de DNase RQ, et le tout est ramené à un volume de 10 µL avec de l'eau DEPC. Une incubation de 30 min à 37 °C permet d'activer l'enzyme DNase qui est ensuite inhibée par ajout de solution stop RQ1 (20 mM EGTA, pH 8) à raison de 1 µL et incubation à 65 °C pendant 10 min.

La transcription inverse a permis d'obtenir des ADN complémentaires (ADNc) à partir des ARN messagers (ARNm) présents dans les ARN totaux. Les ADNc sont ainsi le reflet des séquences exprimées dans les cellules. L'ensemble des ARN totaux traités à la DNase sont ajoutés à 100 µmol d'amorces oligodT modifié (5'- GCTCGAGC(T)<sub>16</sub>VN - 3') et le volume réactionnel est ajusté à 16 µL avec de l'eau DEPC. Une incubation à 70 °C pendant 5 min permet de dénaturer les ARN, puis le milieu réactionnel est placé sur la glace. Aux 16 µL précédents sont ajoutés : 5 µL de tampon de réaction 5X (Tris HCl 250 mM pH 8,3 ; KCl 375 mM ; MgCl<sub>2</sub> 15 mM ; DTT 50 mM) ; 1,25 µL de dNTP (10mM), 30 U RNAsin® inhibiteur (Promega), 200 U de MMLV-RT (Promega). Le milieu réactionnel est alors placé 60 min à 42 °C et les ADNc synthétisés sont conservés à -20 °C.

### **5.6. PCR en temps réel (Q-PCR)**

La technique de PCR en temps réel (ou Q-PCR ou PCR quantitative) est une technique qui permet de suivre l'apparition des produits amplifiés en temps réel lors de la phase exponentielle d'amplification. Parmi les différentes possibilités de marquage par fluorescence, le SYBR Green est très couramment utilisé. Il s'agit d'un fluorophore s'intercalant entre les doubles brins néoformés de l'ADN et qui va émettre très peu de fluorescence lorsqu'il est libre dans le milieu. Une fois fixé, le SYBR Green fluoresce fortement à 520 nm lorsqu'il est excité à 497 nm. La fluorescence est mesurée à la fin de chaque étape d'élongation de chaque cycle, l'augmentation de fluorescence traduit donc la formation d'amplicons. Toutefois, le SYBR Green s'intercalant entre tous les doubles brins, il y a un risque de détecter des dimères d'amorces ou une amplification

non spécifique. La spécificité des couples d'amorces utilisés est contrôlée en fin de cycle d'amplification grâce à une courbe de dissociation traduisant la diminution de la fluorescence en fonction de l'augmentation de la température. Sur cette courbe de fusion, les amplifications non-spécifiques et les dimères d'amorces apparaissent aussi.

La quantification d'un échantillon par PCR en temps réel, que ce soit pour le niveau d'expression ou pour le nombre de copies d'un gène, s'effectue en déterminant le cycle seuil ou "cycle threshold" (Ct) à partir duquel le nombre de cycles PCR permet d'obtenir une fluorescence du produit d'amplification distincte du bruit de fond. Ce paramètre est bien souvent fixé par l'appareil lui-même. Le Ct est inversement proportionnel au logarithme de base 2 du nombre de copies initiales dans le milieu réactionnel. Donc, plus le Ct d'un échantillon sera faible et plus son nombre de copies initiales sera élevé. Pendant longtemps cette expérimentation reposait sur deux types de quantification. D'une part, la quantification absolue s'appuyant sur des gammes étalons de dilutions, notamment de plasmides ou d'ADNc, permettant de déterminer le nombre de copies initiales et l'expression du gène cible et d'autre part, la quantification relative permettant de comparer l'expression d'un gène cible par rapport à un gène de référence ou gène de "ménage". Actuellement ces deux types de quantification sont réalisés conjointement.

### ***Quantification du niveau d'expression des transcrits***

Des couples d'amorces ont été déterminés sur les gènes d'actine  $\beta$ , 18S, shsp20 et de MLE de *Amphora acutiuscula* et des amorces spécifiques de *Phaeodactylum tricornutum* ont été utilisées pour rechercher puis quantifier l'expression des gènes RPS, actine 12 et ubiquitine (Siaut et al., 2007 ; mémoire Master 2 : S. Thiriet-Rupert, 2013) (Tab. 4). Le kit *GoTaq* QPCR Master Mix (Promega) a été utilisé. Aux 6,7  $\mu$ L d'ADNc (issus des ARN extraits des cultures stressées) sont ajoutés : 7,55  $\mu$ L de master mix (SYBR<sup>®</sup> Green, dNTPs, MgCl<sub>2</sub> et l'enzyme, la *GoTaq*<sup>®</sup> polymérase), 0,1 (Actine et 18S) à 0,8  $\mu$ M de chaque amorce et 0,15  $\mu$ L de CXR (carboxy-X-rhodamine) pour compenser les variations de fluorescence entre les différents puits pendant la Q-PCR, selon les recommandations du fournisseur.

**Tableau 4 : Tableau récapitulatif des gènes étudiés et leurs amorces pour la Q-PCR.**

Gène	Amorces	Séquence
<i>Ubiquitine</i>	<i>AacutUbifw</i>	CGA CAA CGT CAA GAC CAA GAT TC
	<i>AacutUbirv</i>	CCA TCT TCC AGC TGC TTT CC
<i>RPS</i>	<i>AacutRPSfw</i>	CCG TCG TGG AAC AAC AAA TAG TC
	<i>AacutRPSrv</i>	CTT CGA CAA AGG CAC CAT AGG
<i>Actine 12</i>	<i>AacutActin12fw</i>	CGG AAG CGT ACA AGG AAA GG
	<i>AacutActin12rv</i>	CGA GAG CGT ATG ACG CAA ATC ATG
<i>Actine <math>\beta</math></i>	<i>QPCRAacutActinbF</i>	CTT GCC GGT CGC GAT CT
	<i>QPCRAacutActinbR</i>	CGT TCA GCC GTA GTG GTC AA
<i>18S</i>	<i>QPCRAacut18SF</i>	AAA ATT GGA CCA CAG GGA ACA C
	<i>QPCRAacut18SR</i>	CCA GCA GCC GCG GTA AT
<i>sHsp20</i>	<i>QHSPAmpfw</i>	CAT TAA CGA ACC AAC TGC CGC
	<i>QHSPAmprv</i>	CAC ATC AAA TGT ACC ACC ACC AAG
<i>MLE</i>	<i>AcutimarF1</i>	AAA GTG ATG TTC TTG TGT GC
	<i>AcutimarR1</i>	CCA ACA TAT CCA AAT GGC CAA AT
	<i>AcutimarF2</i>	AAG GTG ATG TTC TTG TGT GC
	<i>AcutimarR2</i>	CCA ACA TGT CCA AAT GGC CAA AT
	<i>AcutimarFdeg</i>	AAR GTG ATG TTC TTG TGT GC
	<i>AcutimarRdeg</i>	CCA ACA TRT CCA AAT GGC CAA AT

Cependant, avant toute analyse avec le SYBR green, il convient de vérifier qu'un seul fragment est amplifié. Pour cela, une courbe dite de dissociation (cycle d'augmentation de température, +0,3 °C toutes les 15 secondes de 60 °C à 95 °C) est effectuée pour chaque échantillon à la fin du programme d'amplification. Si un seul pic est observé, alors un seul fragment est amplifié. Il convient de vérifier que la température de dissociation déterminée est bien spécifique du fragment à amplifier. Ensuite, une courbe standard (logarithme de la concentration de la matrice en fonction des valeurs de Ct) est réalisée afin de déterminer l'efficacité des amorces utilisées. Une amplification est donc réalisée sur une gamme de dilutions d'ADNc (1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128). L'efficacité est calculée grâce à la pente de la droite obtenue et doit être comprise entre 85 % et 115 %. Le R<sup>2</sup> doit, quant à lui, être supérieur à 0,98.

Par la suite, chaque gène sélectionné (gène de ménage ou d'intérêt) a été amplifié sur des ADNc, obtenus à partir de cultures stressées, et dilués au 1/2, excepté le gène de la sous unité ribosomale 18S dont les transcrits sont toujours plus abondants. Dans ce cas, les ADNc ont été dilués au 1/128. Les ADNc dilués sont amplifiés dans les mêmes conditions que précédemment. Les contrôle(s) endogène(s) appelé(s) "gène(s) de ménage" ou gène(s) de référence, permettent de calibrer les données en corrigeant l'effet des variations de la quantité initiale d'ARN ou encore l'efficacité de l'étape de



transcription inverse. L'utilisation d'au moins 3 gènes de ménage pour normaliser l'expression relative est maintenant très fortement recommandée (Vandesompele et al., 2009). Ceux-ci sont sélectionnés en utilisant le logiciel BestKeeper (Pfaffl et al., 2004). Le calcul d'écart type à partir des valeurs de Ct et de coefficients de corrélation entre les gènes de ménage utilisés permet cette sélection. Le calcul des expressions relatives (ER) des gènes d'intérêt par rapport aux gènes de ménage est basé sur la méthode de comparaison des Ct et déterminé par la formule suivante :

$$ER = \frac{(1 + E \text{ du gène d'intérêt})^{Ct \text{ du calibrateur} - Ct \text{ de l'échantillon}}}{(1 + E \text{ du calibrateur})^{Ct \text{ du calibrateur} - Ct \text{ de l'échantillon}}}$$

Lorsque le résultat est supérieur à 1, il indique le niveau de surexpression du gène d'intérêt comparativement au calibrateur et, lorsqu'il est inférieur à un, le rapport 1/ER donne le niveau de sous-expression.

### ***Quantification du nombre de copies de MLE dans le génome d'*A. acutiuscula****

La PCR en temps réel est également une technique permettant de déterminer le nombre de copies d'un gène dans le génome d'un organisme. Elle a été utilisée tant chez les animaux (Ballester et al., 2004), que chez les végétaux (Mason et al., 2002) ou encore chez les champignons (Schnerr et al., 2001) et notamment dans le cas du nombre de copies de transposons (Kolacsek et al., 2011).

Pour cela, la matrice utilisée n'est pas de l'ADNc mais de l'ADN génomique. La fluorescence obtenue lors de l'amplification du gène ciblé étant proportionnelle au nombre de copies de ce gène, celui-ci est ensuite calculable en utilisant un calibrateur. Ce calibrateur peut être un gène présent en une seule copie dans le génome étudié (Shepherd et al., 2009) ou bien une solution contenant un nombre connu de plasmide (Caprioara-Buda et al., 2012). Dans notre cas, ces plasmides portent un insert correspondant à un fragment de 96 pb du gène d'intérêt (transposase de MLE) et la solution de plasmides purifiée est dosée selon la méthode de Whelan (2003). La Q-PCR a été faite sur le plasmide dilué et l'efficacité des amorces *AcutimarF1/R1* (5' - AAAGTGATGTTCTTGTGTG-3' et 5' - CCAACATATCCAAATGGCCAAAT-3') sur ADNg a été déterminée à partir d'une gamme de dilutions (1/40, 1/80, 1/160, 1/320 et 1/640).

Le nombre de copies dans l'échantillon est déterminé ensuite suivant les instructions de Eagle et Crease (2012), c'est-à-dire en utilisant la méthode du  $\Delta C_t$  mais avec la correction par la PAE (Percentile amplification efficiency) de Yuan et al. (2008). Le nombre de copies correspond donc à  $2^{-\Delta C_t}$  où :

$$\Delta C_t = (C_{t_{MLE \text{ sur ADNg}}} \times PAE_{MLE \text{ sur ADNg}}) - (C_{t_{MLE \text{ sur plasmides}}} \times PAE_{MLE \text{ sur plasmides}})$$

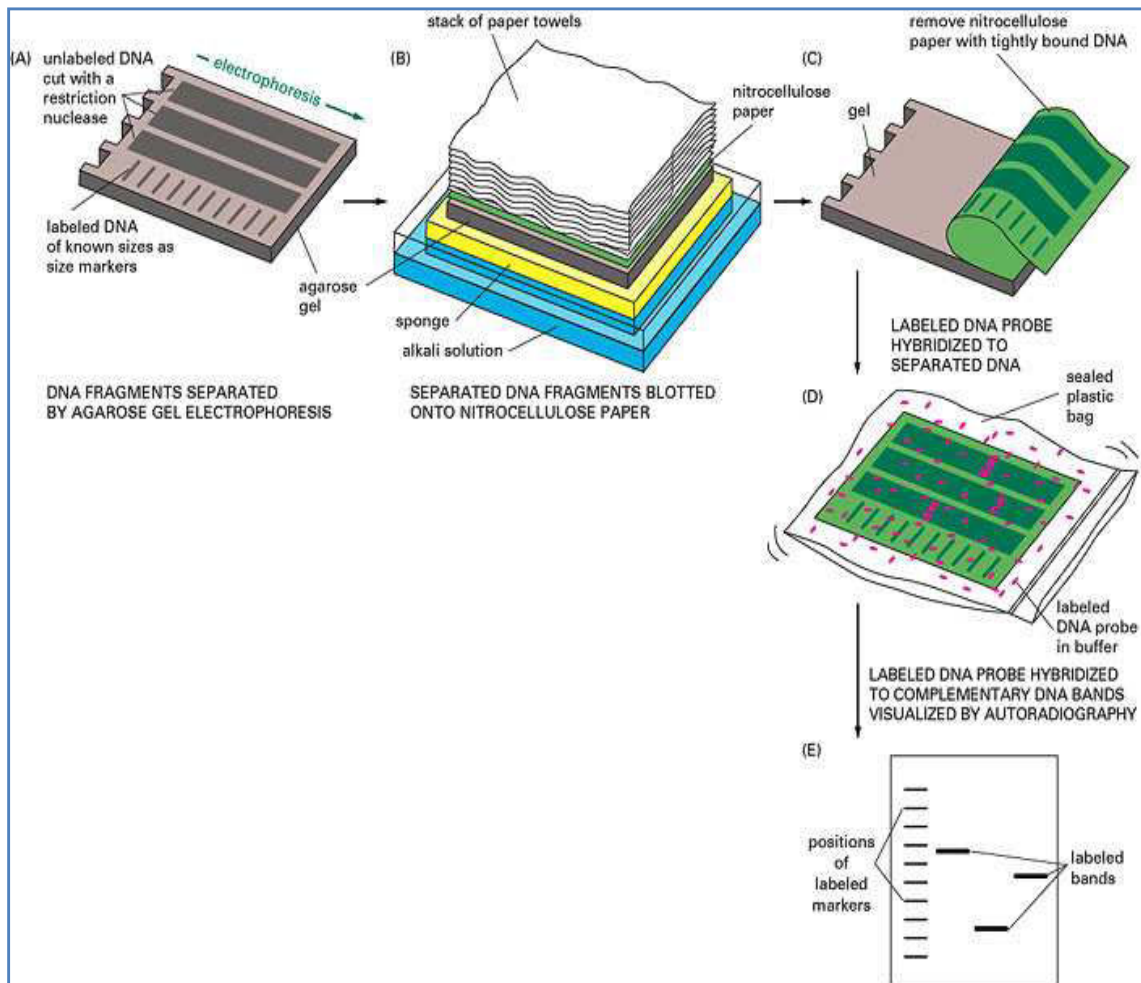
$$PAE = \frac{\ln(AE)}{\ln(2)} \quad \text{et} \quad AE = 10^{-1/\text{pente}}$$

Le nombre obtenu correspond alors au nombre de copies dans l'extrait d'ADN obtenu à partir de nombreuses diatomées. Il faut donc diviser ce résultat par le nombre de cellules utilisées pour l'extraction d'ADN. On obtient alors le nombre de copies de MLE par cellule et donc dans le génome de la diatomée *Amphora acutiuscula* NCC216.

### **5.7. Southern blot**

Huit  $\mu\text{g}$  d'ADNg sont digérés grâce à 40 U d'enzyme HindIII en présence de tampon spécifique 1X de ces enzymes dans 300  $\mu\text{L}$  d'eau stérile. Le mélange est incubé une nuit à 37 °C. Vingt microlitres de ces produits de digestion sont déposés sur un gel d'agarose 1 % afin de vérifier la qualité de la digestion de l'ADN. Aux 280  $\mu\text{L}$  restants, sont ajoutés 3 volumes d'éthanol 100 % et 1/10 de volume d'acétate de sodium 3 M et le tout est incubé 1 h à - 20 °C puis centrifugé 20 min à 4 °C et 12000g. Le culot séché à l'air est repris dans 50  $\mu\text{L}$  de TE puis déposé sur un gel d'agarose 0,8 %. Une migration de 3 h est effectuée à 90 V avant de visualiser la digestion sous UV.

Le gel est rincé 1 min à l'eau ultra pure puis incubé 20 min dans le tampon de dépurination (HCl 0,25 N) afin de faciliter le transfert sur la membrane. Entre deux tampons différents, le gel est lavé 10 min dans l'eau ultra pure. Trois bains de 15 min dans du tampon de dénaturation (NaCl 1,5 M, NaOH 0,5 N) sont effectués. Ce traitement permet d'hydrolyser l'ADN dépuriné et d'obtenir des fragments plus courts. Puis le gel est incubé 3 fois 15 min dans un tampon de neutralisation (NaCl 1,5 M, Tris 1 M pH 7,4). Enfin, le gel, une membrane de nylon (Amersham) et deux papiers Whatman sont incubés 10 min dans le tampon de transfert SSC 10X (NaCl 1,5 M, Tris Sodium Citrate 150 mM). Le transfert de l'ADN sur la membrane (Fig. 26) se fait par capillarité pendant une nuit. L'ADN est ensuite fixé sur la membrane de nylon (Amersham) par cuisson pendant 1h30 minutes à 80 °C.



**Figure 26 : Schéma du principe du transfert d'ADN sur membrane par Southern blot.**

Les sondes radiomarquées au  $[\alpha \text{ P}^{32}] \text{ dATP}$  sont réalisées grâce au kit "Prime a Gene Labelling System" (Promega) à partir de fragments séquencés de 400 pb obtenus avec les amorces MLE (Feschotte et Wessler, 2002). Les sondes sont dénaturées et incubées en présence de Random primer, de radioactivité et de l'enzyme Klenow qui permet la synthèse de nouveaux brins radiomarqués. La préhybridation et l'hybridation de la membrane sont réalisées à 65 °C selon les conditions de Sambrook et al. (1989).

La membrane est placée dans une cassette permettant l'amplification des rayonnements, recouverte par un film photo et placée à -80 °C durant 15 jours. Puis le film est révélé par un passage successif dans un bain révélateur (Kodak), un bain d'eau et enfin un bain de fixateur (Kodak).

## **5.8. Clonage de fragments d'ADN**

### ***Production de bactéries chimio-compétentes***

La souche bactérienne *Escherichia coli* DH5 $\alpha$  est rendue compétente à l'aide du protocole suivant. A partir d'une souche de bactérie, conservée dans le glycérol, une pré-culture est réalisée à 37 °C sous agitation pendant une nuit dans 3 mL de milieu SOB (Tryptone 2%, extrait de levure 0,5%, NaCl 10 mM, KCl 2,5 mM, MgCl<sub>2</sub> 10 mM, MgSO<sub>4</sub> 10 mM). 1 mL de cette pré-culture sert d'inoculum pour ensemer 100 mL de milieu SOB qui est mis à incuber à 37 °C sous agitation jusqu'à l'obtention d'une absorbance égale à 0,5 à 600 nm. Après une incubation de 10 min sur glace, la culture est centrifugée 20 min à 3 000 rpm à 4 °C. Le culot bactérien est mis en suspension dans 20 mL de solution froide pour bactéries compétentes (CaCl<sub>2</sub> 80 mM, MnCl<sub>2</sub> 20 mM, MgCl<sub>2</sub> 10 mM, acétate de potassium 10 mM, glycérol 10%, pH 6,4 ; filtrée sur 0,2  $\mu$ m (Hanahan et Bloom, 1996). Après une incubation de 30 min sur glace, le tout est centrifugé pendant 10 min à 3000 rpm à 4 °C. Le culot bactérien est doucement remis en suspension dans 3 mL de solution froide avec un cône de pipette 1000 coupé sans faire "mousser", puis incubé 10 min sur glace, avant d'être répartis en volumes de 100  $\mu$ L et immédiatement stockés à -80 °C pendant 3 mois (durée maximale de compétence).

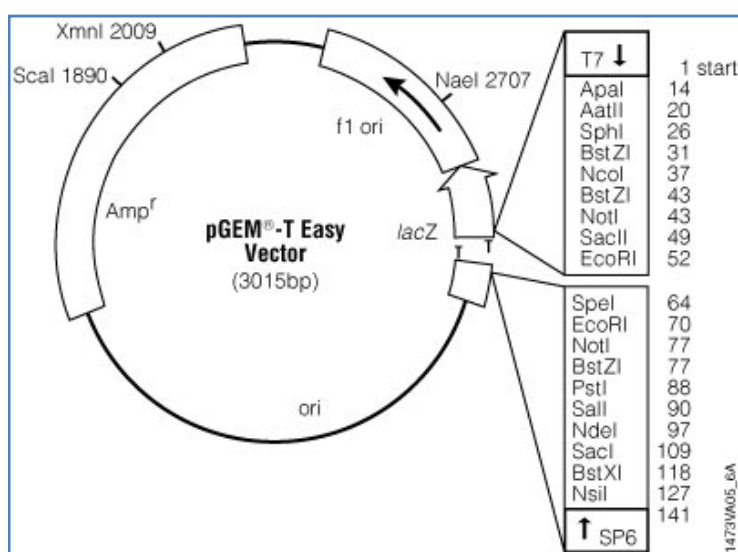
Le contrôle de l'efficacité de la transformation (compétence des bactéries) est réalisé avec le plasmide pUC19 Invitrogene à 10 pg/ $\mu$ L. L'obtention de 10<sup>6</sup> à 10<sup>8</sup> bactéries transformées par  $\mu$ g ADN plasmidique permet de considérer que les bactéries sont bien compétentes.

### ***Élution des fragments d'ADN à partir d'un gel d'agarose***

Les fragments d'ADN sont séparés et visualisés parallèlement aux marqueurs de taille par électrophorèse sur gel d'agarose 1 % et 0,5 % BET (Bromure d'éthidium), ce qui permet de déterminer la taille relative des fragments obtenus. Les amplifications d'ADN sont récupérées en coupant directement le gel, qui pourra être conservé à -20 °C si besoin. L'ADN est ensuite extrait du gel en utilisant le kit d'extraction "Wizard® SV Gel and PCR Clean-Up System" (Promega) selon les instructions du fournisseur. Ce kit permet de purifier des fragments de tailles comprises entre 100 pb et 10 Kb par accrochage sur membrane et centrifugation. Après un lavage rapide, l'ADN est élué à l'aide au moins 30  $\mu$ L d'eau ultra pure et stocké à -20 °C ou prêt pour d'autres applications.

### ***Ligation des produits d'élution dans le vecteur pGEM<sup>®</sup>-T Easy***

Les fragments d'ADN purifiés ont été intégrés dans des plasmides pGEM<sup>®</sup>-T Easy (Promega) (Fig. 27) à l'aide de la T4 DNA ligase. Ce vecteur de clonage de 3015 pb possède un gène de résistance à l'ampicilline pour la sélection des bactéries. Il est fourni sous forme ouverte et additionné de thymidine aux extrémités 3' terminales. Les séquences d'intérêt sont insérées à l'intérieur de son site de clonage situé dans le gène LacZ. Ce vecteur contient de nombreux sites de restriction permettant l'extraction de l'insert par digestion. Il contient aussi les promoteurs pour les ARN polymérase T7 et SP6 et une origine de réplication du phage f1, pour la préparation d'ADN simple brin, et une origine de réplication procaryote.



**Figure 27 : Schéma du vecteur pGEM<sup>®</sup>-T Easy.**

Le protocole de ligation basé sur les instructions du fournisseur est le suivant. Le mélange réactionnel de ligation est composé de 1-3 µL de produit PCR purifié ; tampon de ligation 2X (Tris-HCl 60 mM ; MgCl<sub>2</sub> 20 mM ; DTT 20 mM ; ATP 2 mM et polyéthylène glycol 10% ; pH 7,8), 3 U d'enzyme T4 DNA Ligase , 50 ng vecteur pGEM<sup>®</sup>-T Easy . Le volume réactionnel est finalement complété à 10 µL avec de l'eau ultra pure à la suite d'un mélange par micropipette. La ligation est réalisée pendant une nuit à 4 °C. Lors de l'étape de ligation, le vecteur pGEM<sup>®</sup>-T Easy est refermé.

### ***Transformation et sélection bactérienne***

Les plasmides recombinés sont introduits dans des bactéries chimio-compétentes permettant le clonage du plasmide et donc du fragment d'intérêt par multiplication bactérienne. Le produit de ligation (10  $\mu$ L) est inoculé dans 50  $\mu$ L de suspension de bactéries chimio-compétentes. Une incubation de 30 min sur glace et un choc thermique à 42 °C pendant 45 – 50 s au bain-marie permettent la transformation par intégration des plasmides dans les bactéries. Celles-ci sont ensuite placées durant 2 min sur glace pour restaurer la rigidité de la paroi bactérienne. Les bactéries sont mises en culture dans du milieu de LB (Luria Bertani) liquide (Tryptone 10 g/L ; extrait de levure 5 g/L ; NaCl 5 g/L; pH 7,5) à 37 °C à l'horizontale sous agitation douce à 160 rpm. Après centrifugation pendant 5 min à 900 g, les culots bactériens sont remis en suspension de manière homogène dans le milieu de culture à l'aide de la micropipette puis stérilement étalées, avec des billes de verre, sur les boîtes de Pétri contenant un milieu LB solide auquel sont ajoutés du X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside) à 2 %, de l'IPTG (Isopropyl- $\beta$ -D-1-thiogalactopyranoside) à 20 mg/mL et de l'ampicilline à 100  $\mu$ g/mL. Les boîtes de Pétri sont mises dans l'étuve à 37 °C toute la nuit. Elles sont finalement placées à 4 °C pour stopper la croissance bactérienne tout en préservant leur viabilité.

La présence d'un antibiotique Ampicilline 100  $\mu$ g/mL dans l'agar va empêcher les bactéries sans plasmide (non-transformées) de se multiplier. Les colonies bactériennes ayant intégré le fragment d'intérêt sont discriminées par criblage blanc ou bleu. Les colonies bleues sont non-recombinantes, le gène lacZ est actif et produit l'enzyme  $\beta$ -galactosidase hydrolysant le X-gal avec formation d'un précipité bleu. Au contraire, les colonies blanches sont recombinantes et présentent une inactivation du gène lacZ indiquant que la protéine n'est plus synthétisée.

### ***Extraction d'ADN plasmidique***

Après la mise en culture d'une colonie bactérienne sélectionnée dans 3 mL de LB liquide additionné d'Ampicilline 100  $\mu$ g/mL durant une nuit à 37 °C, la purification d'ADN plasmidique est réalisée à l'aide du kit Wizard®*Plus* SV Minipreps DNA Purification System (Promega). Le protocole d'extraction utilisé est le suivant. Deux mL de culture bactérienne sélectionnée sont centrifugés pendant 5 min à 900 g et à température ambiante, le culot est repris dans 250  $\mu$ L de solution de resuspension. Après avoir doucement agité, 250  $\mu$ L de solution de lyse cellulaire sont ajoutés puis 10  $\mu$ L de solution alcaline de protéase sont additionnés. L'ensemble est homogénéisé au

vortex et incubé pendant 5 min à température ambiante. 350 µL de solution de neutralisation sont ajoutés et une centrifugation de 10 min à la vitesse maximale permet de séparer le surnageant des débris cellulaires. Le surnageant contenant les plasmides clonés est filtré sur mini-colonne. Après filtration, le culot est purifié deux fois grâce à la solution de lavage. Les plasmides sont élués avec 50 µL d'eau ultra pure à température ambiante. La présence de l'insert est vérifiée par PCR et la concentration de plasmides extraits est estimée par mesure de la densité optique 260 nm, puis les plasmides sont envoyés au séquençage ou stockés à -20 °C.

### **5.9. Séquençage des fragments d'intérêt**

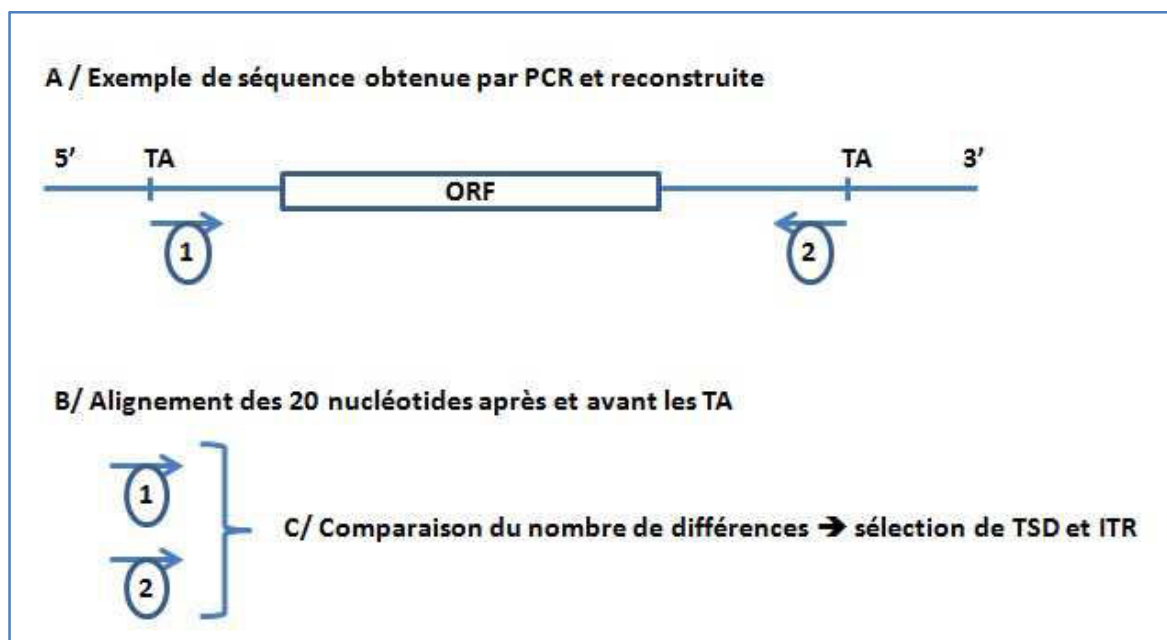
Les plasmides contenant le fragment d'intérêt, ou du moins un fragment à la taille attendue, ont été séquencés par la Société Cogenics Beckman Coulter Genomics (<https://www.cogenicsonline.com>). Les amorces utilisées sont SP6 et T7 qui s'hybrident aux régions bordant l'insert cloné dans le vecteur pGEM-T Easy. Pour des séquences ayant une taille supérieure à 600 pb, quelques amorces supplémentaires sont désignées sur la séquence connue pour avoir de meilleurs séquençages.

### **5.10. Bioinformatique**

Les séquences nucléotidiques ont été comparées entre elles par alignement à l'aide du logiciel de bioinformatique ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Les identifications de séquences, notamment les gènes MLE, ont été réalisées avec NCBI BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Après la traduction de la séquence nucléotidique en séquences d'acides aminés, des motifs HTH ont été mis en évidence en utilisant un logiciel de prédiction de motifs : [http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_hth.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hth.html). Des structures du signal de localisation nucléaire (SLN) sont prédites par le logiciel NLS Mapper ([http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS\\_Mapper\\_form.cgi](http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi)). La boîte TATA est identifiée par le logiciel HCtata: Hamming-Clustering Method for TATA Signal Prediction in Eukaryotic Genes ([http://zeus2.itb.cnr.it/~webgene/wwwHC\\_tata.html](http://zeus2.itb.cnr.it/~webgene/wwwHC_tata.html)). Le pourcentage en GC dans les MLE et les ORF sont comptés sur le site : <http://www.encorbio.com/protocols/Nuc-MW.htm>. Le pourcentage en GC en troisième position dans l'ORF est réalisé avec le logiciel statistique R.

Afin de détecter des transposons complets, les ITR ont été recherchés dans les séquences reconstruites. Cette recherche a été réalisée en collaboration avec J-D

Rouault du laboratoire Évolution, Génomes et Spéciation IDEEV CNRS A9034 de Gif-sur-Yvette. JD Rouault a développé un algorithme de recherche des ITR en se basant sur la localisation de TSD potentiels. Pour cela, toutes les occurrences de dinucléotide TA (TSD pour MLE) situées dans une zone comprise entre 0 et 1000 pb en amont du gène de la transposase ont été comparées à une zone de taille identique située après l'ORF. Ainsi, ont été alignées toutes les séquences de 20 nucléotides localisées juste derrière un premier TA avec celles de 20 nucléotides se trouvant immédiatement en avant d'un second TA. Pour chaque couple de séquences répertoriées, le nombre de différences a été compté. Si ce nombre est inférieur à un seuil fixe prédéfini, par exemple 0, 2 ou 4, les deux TA et leurs séquences associées ont été extraites et considérés comme des TSD et des ITR potentiels (Fig. 28). Le passage du statut de ITR potentiel à définitif est empirique et statistique, quand on a trouvé plusieurs couples de ITR identiques ou très proches encadrant des ORFs quasi identiques.



**Figure 28 : Schéma de la démarche informatique simplifiée pour détecter les ITR.**

A) Tous les dinucléotides TA situés entre 0 et 1000 pb en amont et en aval du gène de la transposase ont été recensés. B) Toutes les séquences de 20 nucléotides situées après le TA (flèche 1) et celles situées avant le TA (flèche 2) ont été recensées.



## **6. Phylogénie**

### **6.1. Matériel moléculaire**

Cinquante cinq séquences de gènes complets codant une transposase appartenant à la superfamille *Tc1-mariner* ont été collectées à partir de la banque de données GenBank. Parmi la famille *Tc1*, nous avons choisi des représentants bien identifiés tels que *Minos*, *impala*, *paris*, *Topi*, *tana1*, *tana2* et *Tc1* qui ont tous une structure de type DD34E. Dans la famille *mariner*, au moins deux représentants de chaque sous-famille ont été ajoutés au jeu de données qui se compose des groupes suivants DD34D, DD37D, DD39D, DD41D, DD45D (Tab. 5). Dans le groupe DD34D, deux à trois éléments ont été sélectionnés dans chaque infragroupe: *elegans*, *cecropia*, *mellifera*, *mauritiana*, *marmoratus*, *irritans*. Pour les éléments DD39D ou "plant *mariner*", qui apparaissent comme les plus proches des séquences MLE d'*A. acutiusculad*'après les alignements avec les bases de données, toutes les séquences complètes de transposases connues ont été utilisées.

### **6.2. Méthode de phylogénie**

Toutes les séquences du jeu de données créé pour l'analyse phylogénétique ont été traduites en acides aminés et alignées à l'aide du logiciel ClustalW2 sur le site EMBL-EBI (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). L'alignement multiple qui en résulte a été amélioré manuellement dans le logiciel BioEdit version 7.1 (Hall, 1999) afin d'optimiser l'alignement sur la base des motifs conservés tels que ceux de la triade catalytique. L'alignement de 744 acides aminés produits a été phylogénétiquement analysé par Inférence Bayésienne (IB) qui correspond à la meilleure méthode permettant d'analyser le jeu de données contenant des séquences présentant des structures différentes. Le modèle WAG+Ga a été sélectionné pour expliquer la distribution des données. Ce modèle est le meilleur selon le critère AIC par ProtTest version 2.4 ([http://darwin.uvigo.es/software/prottest2\\_server.html](http://darwin.uvigo.es/software/prottest2_server.html)). Ensuite, le modèle WAG+G a été implémenté dans le logiciel MrBayes 3.1 (Ronquist et Huelsenbeck 2003) utilisant l'inférence bayésienne. L'analyse phylogénétique et la construction de l'arbre ont été réalisées trois fois afin de vérifier la cohérence des arbres produits. Pour chaque analyse, deux séries (runs) ont été effectuées parallèlement, en vérifiant la valeur du PSRF qui, lorsqu'il est proche de 1, indique que les deux séries sont convergentes après 500 000 générations. Après élimination de 25% du bruit de fond, nous avons présenté l'arbre consensus en indiquant les valeurs de probabilité postérieure supérieures à 50% grâce au logiciel MEGA4 version 4.0 (Tamura et al. 2007).

**Tableau 5 : Tableau récapitulatif des ET complets de la superfamille *Tc1-mariner* et *ITm* utilisés pour l'étude phylogénétique.**

	Elément	Numéro d'accession	Triade catalytique	Espèce	Référence	
<i>ITm</i>	<i>Ag1</i>	AF378002	DD37E	<i>Anopheles gambiae</i>	Shao et Tu, 2001	
	<i>Aa1</i>	AF377999	DD37E	<i>Ochlerotatus atropalpus</i>	Shao et Tu, 2001	
<i>Tc1-like</i>	<i>Tc1</i>	P03934	DD34E	<i>Caenorhabditis elegans</i>	Rosenzweig et al., 1983	
	<i>Paris</i>	AAA88882	DD34E	<i>Drosophila virilis</i>	Petrov et al., 1995	
	<i>Minos</i>	Q24397	DD34E	<i>Drosophila hydei</i>	FranzetSavakis, 1991	
	<i>Topi1</i>	AAD03792	DD34E	<i>Anopheles gambiae</i>	Grossman et al., 1999	
	<i>Impala</i>	AAB33090	DD34E	<i>Fusarium oxysporum f. sp. melonis</i>	Langin et al., 1995	
	<i>Tana1</i>	JX889426	DD38E	<i>Acipenser naccarii</i>	Pujolar et al., 2013	
	<i>Tana2</i>	JX889427	DD38E	<i>Acipenser naccarii</i>	Pujolar et al., 2013	
	<i>IE chez animaux</i>	<i>mauritiana</i>	<i>Mos1</i>	X78906	DD34D	<i>Drosophila mauritiana</i>
<i>Mbmar1</i>			AF465247	DD34D	<i>Mamestra brassicae</i>	Mandrioli, 2003
<i>Dteismar1</i>			AAC28261	DD34D	<i>Drosophila teissieri</i>	Brunet et al., 1996
<i>irritans</i>		<i>Himar2</i>	U11641	DD34D	<i>Haematobia irritans</i>	Robertson et Lampe, 1995
		<i>Mpmar1</i>	U11649	DD34D	<i>Mantispa pulchella</i>	Robertson et Lampe, 1995
		<i>Cpmar1</i>	L06041	DD34D	<i>Chrysoperla plorabunda</i>	Robertson et al., 1992
		<i>Hsmar2</i>	U49974	DD34D	<i>Homo sapiens</i>	Robertson et Martos, 1997
<i>mellifera</i>		<i>Ammar1</i>	AAO12861	DD34D	<i>Apis mellifera</i>	Unpublished
		<i>Ccmar2</i>	AY155493	DD34D	<i>Ceratitis capitata</i>	Gomulski et al., 1997
		<i>Ccmar1</i>	AAB17945	DD34D	<i>Ceratitis capitata</i>	Gomulski et al., 1997
		<i>Famar1</i>	AY155492	DD34D	<i>Forficula auricularia</i>	Lampe et al., 2003
		<i>Cemar1</i>	BAB86288	DD34D	<i>Apis cerana</i>	Sumitani et al., 2002
<i>cecropia</i>		<i>Fsmar1</i>	BAB32436	DD34D	<i>Fungia sp. Kusabiraishi</i>	Nakajima et al., 2002
		<i>Dtmar1</i>	CAA50801	DD34D	<i>Girardia tigrina</i>	Garcia-Fernàndez et al., 1993
		<i>Bmmar2</i>	BAA23532	DD34D	<i>Bombyx mori</i>	Tomita et al., 1997
		<i>Hsmar1</i>	U52077	DD34D	<i>Homo sapiens</i>	Robertson et Zumpano, 1997

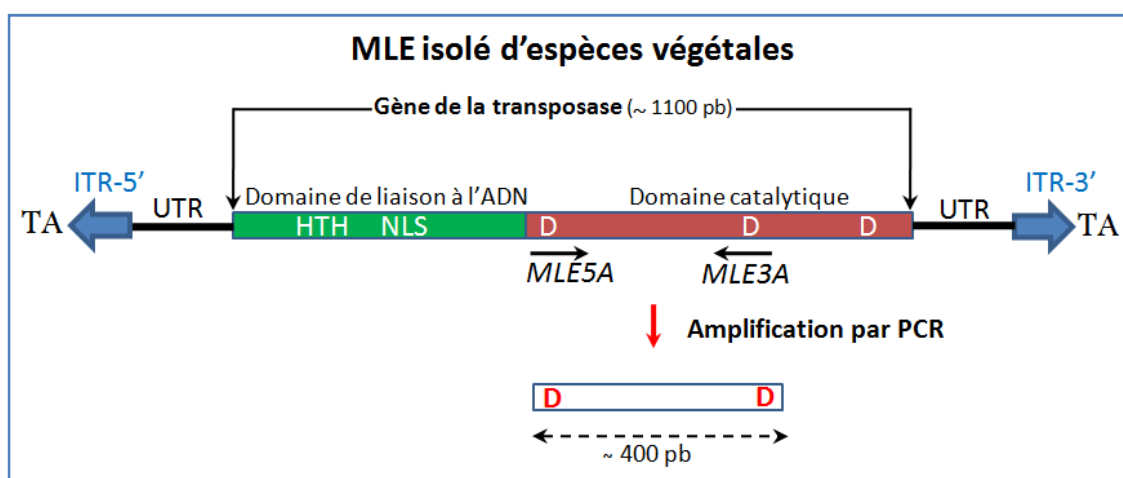
<i>MLE chez les plantes</i>	<i>Pacmmarcon</i>	CAJ76984	DD34D	<i>Pachygrapsus marmoratus</i>	Bui et al., 2008
	<i>Porpmarcons</i>	CAP20050	DD34D	<i>Portunus pelagicus</i>	Bui et al., 2008
	<i>MhmaT1</i>	AEO90418	DD37D	<i>Misgolas hubbardi</i>	Unpublished
	<i>Bmmar1</i>	AAB47739	DD37D	<i>Bombyx mori</i>	Robertson et Asplund, 1996
	<i>Cemar1</i>	AC024800	DD34D	<i>Caenorhabditis elegans</i>	<i>C. elegans</i> Sequencing Consortium
	<i>Cemar5</i>	AF024498	DD34D	<i>Caenorhabditis elegans</i>	<i>C. elegans</i> Sequencing Consortium
	<i>Crosa6</i>	AAK61417	DD41D	<i>Ceratitis rosa</i>	Gomulski et al., 2001
	<i>Tcp32</i>	Z95154	DD41D	<i>Cydia pomonella granulovirus</i>	Jehle et al., 1998
	<i>Cbl</i>	HM101510	DD39D	<i>Chimonobambusa marmorea</i>	Zhou et al., 2011
	<i>Mb469</i>	HM101514	DD39D	<i>Melocanna baccifera</i>	Zhou et al., 2011
	<i>Ba469</i>	HM101501	DD39D	<i>Bambusa bambos</i>	Zhou et al., 2011
	<i>Ss460</i>	HM101490	DD39D	<i>Sasa sinica</i>	Zhou et al., 2011
	<i>Oa499</i>	HM101496	DD39D	<i>Otatea acuminata</i>	Zhou et al., 2011
	<i>Ca499</i>	HM101511	DD39D	<i>Chusquea coronalis</i>	Zhou et al., 2011
	<i>Osmar1</i>	AF172282	DD39D	<i>Oryza sativa</i>	Tarchini et al. 2000
	<i>Psmar1A</i>	AY833549	DD39D	<i>Pisum sativum</i>	Macas et al., 2005
	<i>Soymar1</i>	AF078934	DD39D	<i>Glycine max</i>	Jarvick et Lark, 1998
	<i>mth2</i>	AC137603	DD39D	<i>Medicago truncatula</i>	Unpublished
	<i>Vulmar1</i>	AJ556159	DD39D	<i>Beta vulgaris</i>	Jacobs et al. 2004
	<i>Atmar1</i>	AF262043	DD39D	<i>Arabidopsis thaliana</i>	The <i>A. thaliana</i> Genome Sequencing
<i>mariner chez les protistes</i>	<i>Costa</i>	JX013907	DD39D	<i>Pseudo-nitzschia multiseriis</i>	Unpublished
	<i>B10</i>		DD39D	<i>Phytophthora infestans</i>	Supercontig 1.10 from the genome available on <a href="http://www.broadinstitute.org/annotation/genome/phytophthora_infestans/MultiDownloads.html">http://www.broadinstitute.org/annotation/genome/phytophthora_infestans/MultiDownloads.html</a>
	<i>B21</i>		DD39D	<i>Phytophthora infestans</i>	
	<i>Ot1</i>	CAL52166.2	DD45D	<i>Ostreococcus tauri</i>	Derelle et al., 2006
	<i>Ot2</i>	XP003074897	DD45D	<i>Ostreococcus tauri</i>	Derelle et al., 2006
	<i>Tymar1</i>	AY282463	DD34D	<i>Trichomonas vaginalis</i>	Silva et al., 2005
	<i>marinarEd</i>	EU099441	DD33D	<i>Entamoeba dispar</i>	Lorenzi et al., 2008
	<i>marinerEh</i>	EU099443	DD33D	<i>Entamoeba histolytica</i>	Lorenzi et al., 2008

## PARTIE III : RESULTATS ET DISCUSSION

### Chapitre 1 : Caractérisation des MLE chez *Amphora acutiuscula*

#### 1. Détection de séquences partielles deMLE dans le génome de *A. acutiuscula*

A partir de l'ADN génomique d'*A. acutiuscula*, les éléments transposables de type *mariner* ont été recherchés en utilisant les amorces dégénérées *MLE5A* et *MLE3A* définies par Feschotte et Wessler (2002). Chez les végétaux, ces amorces s'hybrident au niveau des deux premiers D de la triade catalytique, permettant d'amplifier une région d'environ 400 pb localisée dans le gène codant la transposase (Fig. 29).



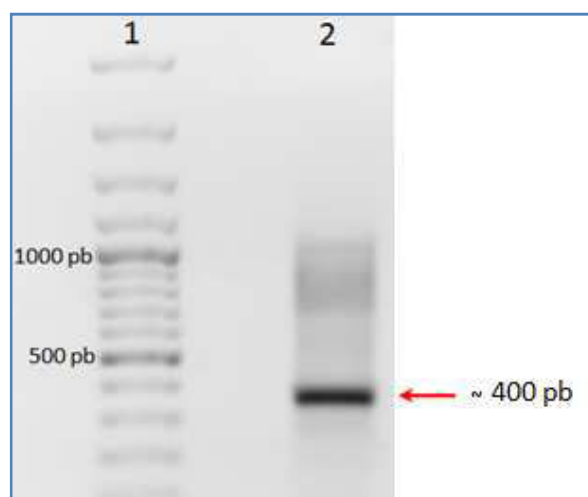
**Figure 29 : Schéma général d'un MLE isolé à partir d'un ADNg végétal, sur lequel sont localisées les amorces *MLE5A* et *MLE3A* ayant servi à la détection des séquences chez *A. acutiuscula*.**

La séquence du gène de la transposase est divisée en 2 parties correspondant, dans la protéine, à son domaine de liaison à l'ADN (ici en vert) et à son domaine catalytique (en rouge). Dans la zone catalytique, les 3 acides aspartiques (D) de la triade ont été positionnés. Les amorces utilisées pour la détection des MLE sont situées au niveau des deux premiers D et appelées *MLE5A* et *MLE3A* (Feschotte et Wessler, 2002). Dans les espèces végétales chez lesquelles elles ont été utilisées, ces amorces ont généralement permis l'amplification de fragments d'environ 400 pb. TA : site cible, dupliqué lors de l'insertion et présent à chaque extrémité de l'élément ; ITR : Répétition Terminale Inversée ; UTR : Région non Traduite.

Chez *A. acutiuscula*, les amorces *MLE5A* et *MLE3A* ont permis l'amplification d'une bande d'environ 400 pb (Fig. 30) correspondant à la taille attendue pour des éléments de type *mariner* isolés de végétaux. Les produits de PCR ont été élués puis clonés dans le plasmide pGEM-T. Deux clones ont été séquencés pour la recherche d'identification des MLE.

Après comparaison de leurs séquences par BlastX avec les banques de données, les deux clones ont révélé une homologie d'environ 50 % avec des transposases de type *mariner* de

plantestelles que l'avoine, le petit pois ou encore le bambou. Chez *A. acutiuscula*, les clones ont été nommés respectivement *Aacutmar-sp 1.7* et *Aacutmar-sp 1.8* (*Aacutmar* pour *Amphora acutiuscula mariner* séquence partielle 1. numéro du clone 7 ou 8). Ils complètent une collection de 6 séquences précédemment obtenues au laboratoire (thèse D. Hermann, 2011) et nommées *Aacutmar-sp 1.1* à *1.6* et déposées dans GenBank sous les numéros FN546185 à FN546190.



**Figure 30 : Électrophorèse sur gel d'agarose 1 % des produits d'amplification obtenus avec les amorces dégénérées MLE5A et MLE3A chez la diatomée *A. acutiuscula*.**

Puits 1 : Marqueur de taille Generuler 100 pb, puits 2 : produits amplifiés. La flèche indique la bande obtenue d'une taille légèrement inférieure à 400 pb qui a été clonée puis séquencée.

Les séquences *Aacutmar-sp 1.7* et *1.8* ont respectivement une taille de 374 et 372 pb (Tab. 6). La comparaison deux à deux des clones *Aacutmar-sp 1.7* et *1.8* avec les éléments *Aacutmar-sp 1.1-1.6* indique une identité comprise entre 81 et 97 % pour *Aacutmar-sp 1.7* et de 77 à 93 % pour *Aacutmar-sp 1.8*. Entre eux, ces éléments montrent une identité de 90 % (Tab. 7). Après traduction en protéines, *Aacutmar-sp 1.7* et *1.8* ont respectivement une longueur de 124 et 123 acides aminés et ne possèdent ni codon stop ni décalage du cadre de lecture du gène de la transposase (Fig. 31). Au niveau protéique, la comparaison de *Aacutmar-sp 1.7* et *1.8* entre eux donne un taux d'homologie de 95 % et leur comparaison deux à deux avec les autres éléments (*Aacutmar-sp 1.1* à *-1.6*) montre des taux d'homologie variant de 79 à 96 % (Tab. 7). Un consensus nommé *ConsAacutmar1-sp* a été établi à partir des huit fragments internes de 380 pb du gène de la transposase. C'est à partir de ce consensus que plusieurs jeux d'amorces ont pu être définis pour la PCR inverse. Cette technique a été utilisée pour rechercher la séquence complète des transposons de type *mariner* dans le génome d'*A. acutiuscula*.

**Tableau 6 : Tableau récapitulant les différentes séquences partielles de MLE obtenues chez la diatomée *A. acutiuscula*.**

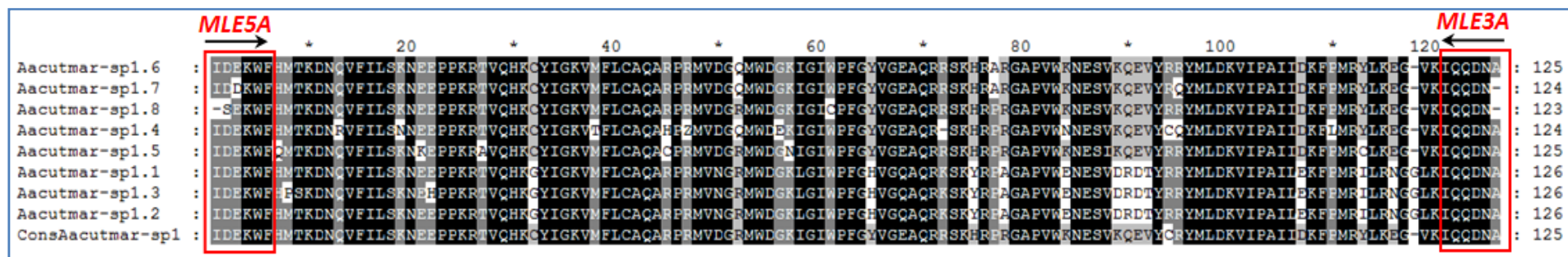
Présent travail et thèse de D. Hermann (2011). Pour chaque séquence, sont précisés : son nom, sa longueur précise, la présence ou non de codons stop et son numéro d'accèsion dans les banques de données.

Nom	Longueur (pb)	Longueur en protéine	ORF ininterrompue	Numéro d'accèsion
<i>Aacutmar-sp1.1</i>	379	126	1	FN546185
<i>Aacutmar-sp1.2</i>	377	125	1	FN546186
<i>Aacutmar-sp1.3</i>	378	126	1	FN546187
<i>Aacutmar-sp1.4</i>	375	124	0	FN546188
<i>Aacutmar-sp1.5</i>	375	125	1	FN546189
<i>Aacutmar-sp1.6</i>	375	125	1	FN546190
<i>Aacutmar-sp1.7</i>	374	124	1	
<i>Aacutmar-sp1.8</i>	372	123	1	

**Tableau 7 : Comparaison deux à deux des huit séquences d'*Aacutmar-sp1* disponibles au laboratoire.**

Les données sont présentées en pourcentages d'homologie nucléotidique et protéique (entre parenthèses pour cette dernière).

Nom	<i>Aacutmar-sp1.1</i>	<i>Aacutmar-sp1.2</i>	<i>Aacutmar-sp1.3</i>	<i>Aacutmar-sp1.4</i>	<i>Aacutmar-sp1.5</i>	<i>Aacutmar-sp1.6</i>	<i>Aacutmar-sp1.7</i>
<i>Aacutmar-sp1.2</i>	99 (100)						
<i>Aacutmar-sp1.3</i>	97 (97)	97 (97)					
<i>Aacutmar-sp1.4</i>	76 (78)	77 (78)	75 (75)				
<i>Aacutmar-sp1.5</i>	79 (80)	79 (80)	78 (77)	92 (86)			
<i>Aacutmar-sp1.6</i>	81 (83)	81 (83)	79 (80)	90 (91)	91 (92)		
<i>Aacutmar-sp1.7</i>	81 (76)	81 (76)	79 (74)	90 (91)	90 (91)	97 (98)	
<i>Aacutmar-sp1.8</i>	79 (78)	78 (78)	77 (75)	93 (88)	92 (92)	92 (96)	90 (95)



**Figure 31 : Alignement des séquences partielles *Aacutmar-sp1.1* à *1.8* traduites en protéines.**

Les amorces utilisées pour l'amplification, *MLE5A* et *MLE3A* (Feschotte et Wessler, 2002), sont encadrées aux extrémités des séquences. Pour la séquence *Aacutmar-sp 1.4*, la phase a dû être restaurée pour permettre l'alignement avec les autres séquences partielles.

## **2. Allongement des séquences de MLE par PCR inverse**

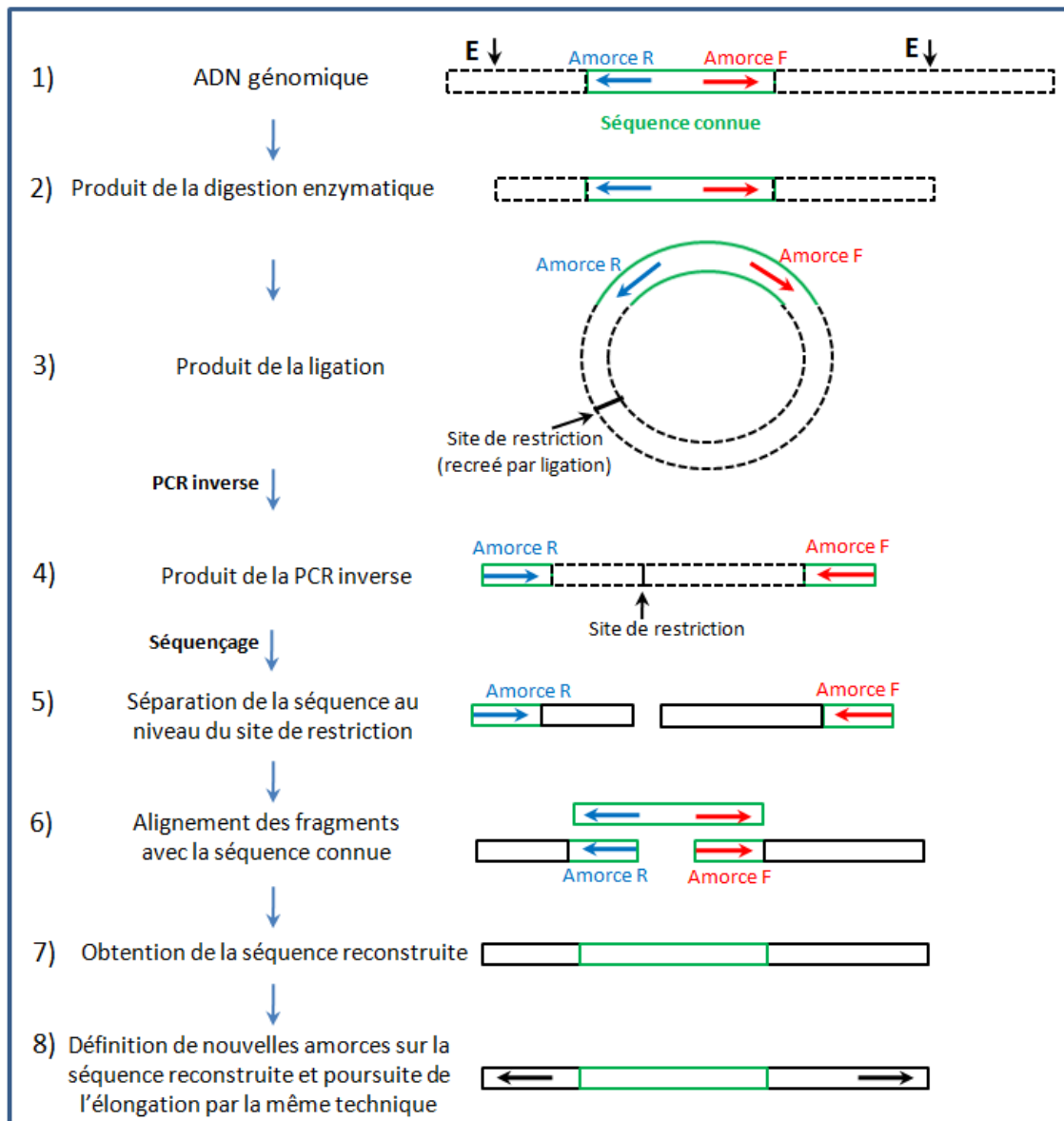
### **2.1. Séquences obtenues par PCR inverse**

La PCR inverse (iPCR) a pour but de connaître les régions inconnues bordant une séquence d'ADN connue. Dans le présent travail, la séquence consensus de 380 pb, *ConsAacutmar-spl*, a servi de point de départ pour amplifier les régions inconnues du transposon avec pour objectif de caractériser un élément complet. La technique d'iPCR a été répétée à plusieurs reprises en fonction de l'avancée le long de l'élément transposable. Pour obtenir un élément complet, l'ADN génomique de la diatomée *A. acutiuscula* a été digéré successivement avec différentes enzymes de restriction et plusieurs couples d'amorces chevauchantes ont été utilisés pour les amplifications. Chaque couple a été défini à partir de la séquence précédemment obtenue, séquence servant à son tour de point de départ pour un nouvel allongement. Ainsi, il a été possible de "marcher" de proche en proche le long du transposon et d'allonger progressivement la séquence connue.

Pour chaque iPCR, les produits d'amplification ont été clonés, séquencés, puis analysés. Les fragments ont été assemblés et ont permis de générer successivement des séquences reconstruites nommées *Re-X* (*Re* pour Reconstituer, *X* n° de la reconstruction, au total ici 4) (Fig. 24), chacune d'elles servant à son tour de base pour un nouvel allongement. Le schéma général récapitulant les étapes de chaque iPCR est présenté dans la figure 23. Ces mêmes schémas présentés dans la partie Matériel et Méthodes ont été déplacés ici pour s'y reporter plus aisément et ainsi faciliter la lecture.

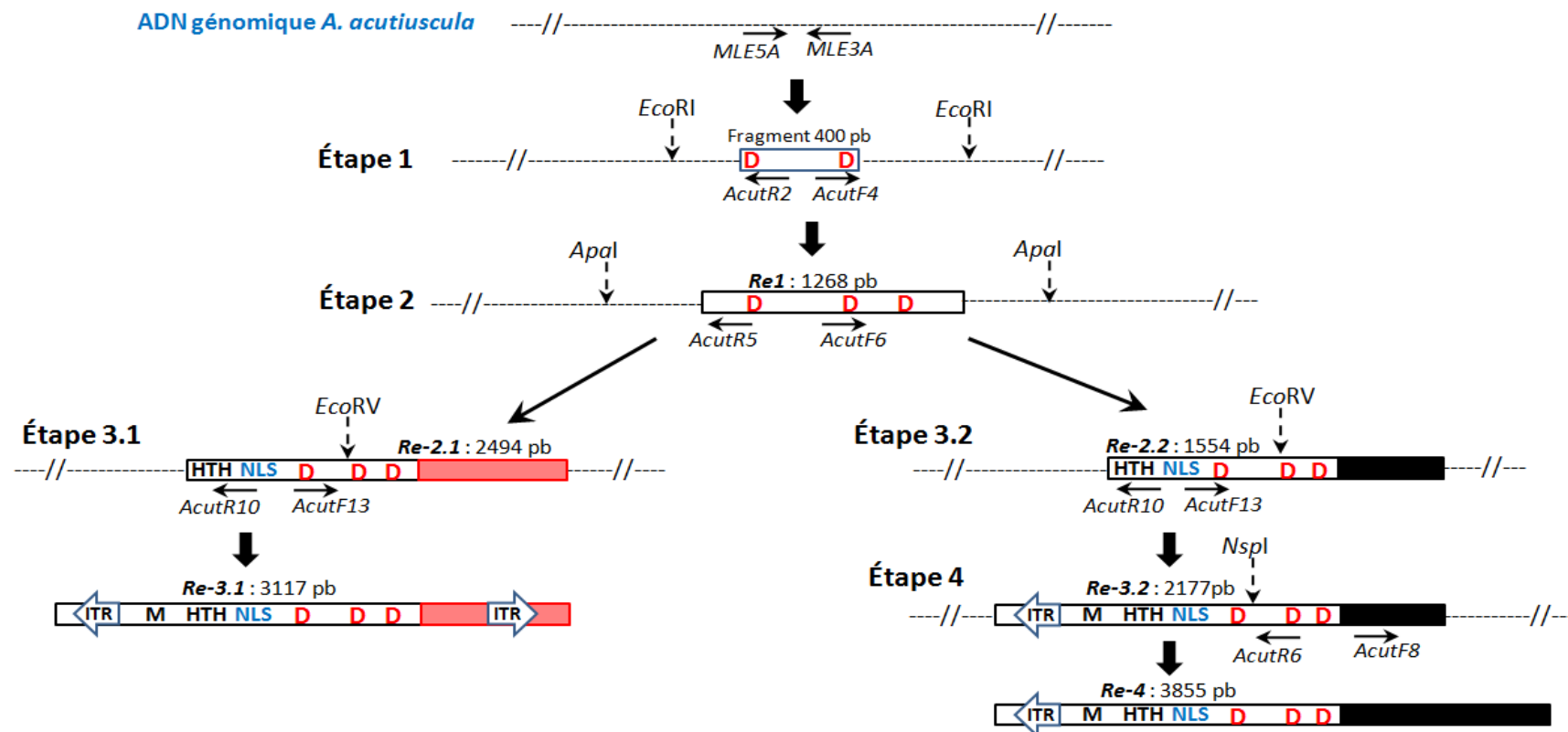
Les étapes menant à l'obtention de différentes séquences reconstruites et qui ont permis d'obtenir un transposon complet de type *mariner* chez *A. acutiuscula* sont décrites ci-après.





**Figure 23 : Schéma récapitulant les étapes de l'allongement d'une séquence connue par PCR inverse.** (Cette figure est aussi présente dans la partie Matériel et méthodes).

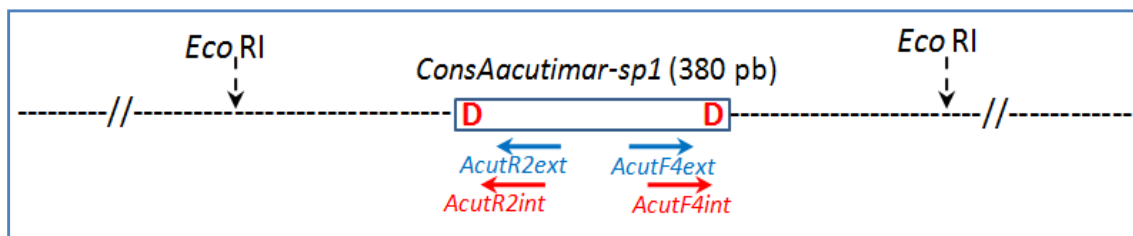
1) et 2) l'ADN génomique de la diatomée *A. acutiuscula* est digéré à l'aide d'une enzyme de restriction qui coupe, ici, à l'extérieur de la séquence connue. 3) Les produits de digestion sont ligués sur eux-mêmes. 4) ils servent de matrice pour la PCR inverse qui utilise deux couples d'amorces chevauchantes, ici seul le second couple est présenté. 5) après séquençage, l'analyse des séquences permet de localiser les amorces et le site de restriction. C'est au niveau de ce site que la séquence est séparée en deux fragments. 6) et 7) Chaque fragment est alors aligné avec la séquence connue, la partie commune entre les fragments et la séquence connue permet de vérifier que c'est bien cette dernière qui a servi de matrice à allongement de la séquence. La partie située au-delà de la séquence connue correspond à l'allongement de la séquence. Une fois les trois fragments assemblés, on obtient une séquence reconstruite. 8) la séquence reconstruite est utilisée comme matrice pour définir de nouvelles amorces qui serviront à poursuivre l'allongement de la séquence en utilisant la même technique. E↓: localisation de la coupure de la digestion enzymatique. Les séquences connues sont indiquées en traits pleins. Les séquences inconnues sont indiquées en pointillés.



**Figure 24 : Schéma récapitulatif des étapes de la reconstruction des séquences MLE chez *A. acutiuscula*.**

Le fragment partiel de MLE (400 pb) a été détecté par l'utilisation des amorces dégénérées *MLE5A* et *MLE3A*. **Étape 1** : l'ADNg d'*A. acutiuscula* est digéré par l'enzyme *EcoRI* qui ne coupe pas le fragment connu de 400 pb. Les amorces *AcutR2/AcutF4* sont utilisées pour construire un fragment *Re-1* de 1268 pb qui contient également le 3<sup>ème</sup> D de la triade catalytique. *Re-1* est le point de départ pour l'iPCR de l'étape 2. **Étape 2** : l'ADNg d'*A. acutiuscula* est digéré par l'enzyme *ApaI* qui ne coupe pas le fragment reconstruit *Re-1*. Les amorces *AcutR5/AcutF6* sont utilisées pour reconstruire deux fragments *Re-2.1* (2494 pb) et *Re-2.2* (1554 bp). Ces fragments partagent une partie identique en région 5' contenant les motifs caractéristiques des MLE et une partie différente en région 3'. **Étape 3** : l'ADNg d'*A. acutiuscula* est digéré par l'enzyme *EcoRV* qui coupe à l'intérieur des *Re-2*. Les amorces *AcutR10/AcutF13* sont utilisées pour allonger leur extrémité 5' et permettre de reconstruire les fragments *Re-3.1* et *Re-3.2*. Nous avons réussi à identifier un transposon MLE complet (*Aacutmar1-re*) dans la séquence de *Re-3.1* avec ses ITR à chaque extrémité. **Étape 4** : l'élongation d'extrémité 3' du *Re-3.2* par l'enzyme *NspI* et les amorces *AcutR6/AcutF8*, cela a permis d'obtenir le *Re-4* (3855 pb).

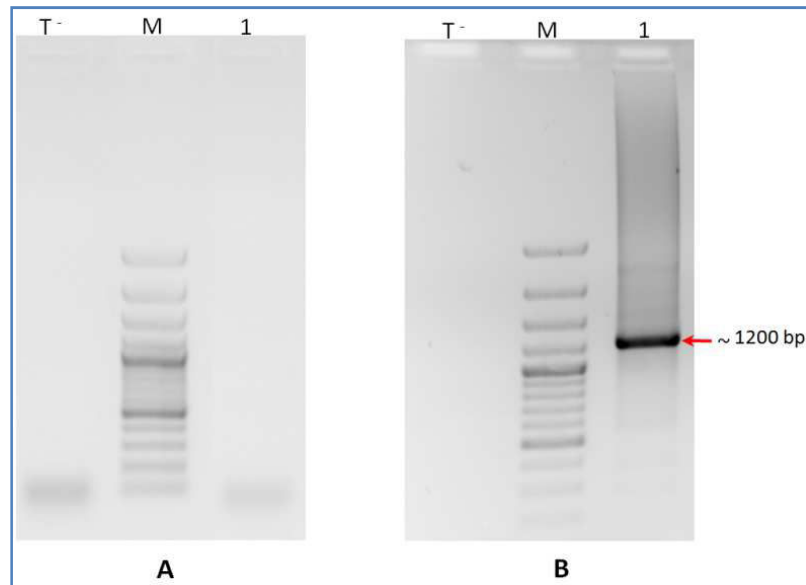
Pour obtenir les régions bordant le fragment de 380 pb, l'ADNg a d'abord été digéré avec l'enzyme *EcoRI* (Fig. 32) qui ne coupe pas dans les séquences *ConsAacutimar-sp1*.



**Figure 32 : Schéma permettant de localiser les amorces ayant servi à la recherche des MLE complets chez la diatomée *A. acutiuscula*.**

Les flèches en bleu et en rouge représentent les deux couples d'amorces utilisés pour la première iPCR sortant du fragment de 380 pb. D'un point de vue technique, chaque iPCR est constituée de deux amplifications successives, les amorces *AcutR2ext* et *AcutF4ext* (pour externe) ont été utilisées pour la iPCR n°1 et les amorces *AcutR2int* et *AcutF4int* (pour interne) pour la iPCR n°2. La ligne pointillée représente la région bordant le fragment de 380 bp.

La première amplification a été effectuée avec le couple d'amorces *AcutR2ext* et *AcutF4ext*. Le dépôt sur gel d'agarose du produit de cette première iPCR ne montre aucune bande visible (Fig. 33-A). La seconde iPCR a été réalisée à partir du produit d'amplification de la première et avec le couple d'amorces *AcutR4int* et *AcutF4int* situé à l'intérieur du premier produit de PCR. A l'issue des deux iPCR successives, un fragment d'environ 1200 pb a été observé (Fig. 33-B).

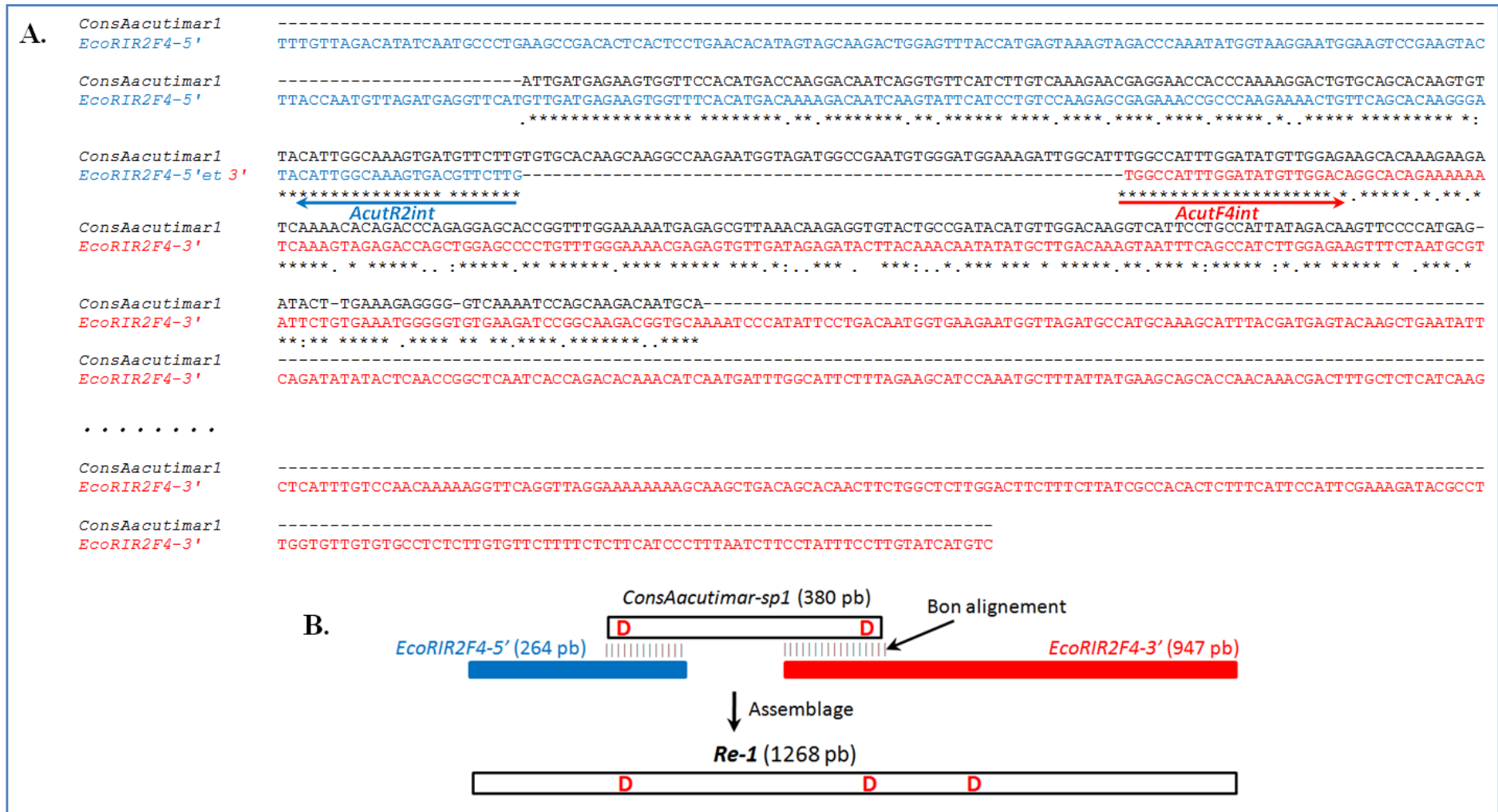


**Figure 33 : Migration sur gel d'agarose des produits d'amplification pour l'obtention des fragments flanquant la séquence *ConsAacutmar-sp1*.**

**A-** Puits 1 : produit de la première iPCR utilisant comme matrice l'ADNg digéré par l'enzyme *EcoRI* et le couple d'amorces externes *AcutR2ext* et *AcutF4ext*. **B -** Puits 1 : produit de la seconde iPCR utilisant comme matrice le produit de l'iPCR précédente et le couple d'amorces internes *AcutR2int* et *AcutF4int*. M : marqueur de taille Generuler 100 pb. T<sup>-</sup> : témoin négatif (iPCR réalisée sans ADNg).

Ce produit a été cloné et deux clones ont été séquencés. Ceux-ci, nommés *EcoRIR2F4-1* et *EcoRIR2F4-2*, ont une longueur exacte de 1217 pb et sont similaires à plus de 99 %, de ce fait, un seul des deux clones a été utilisé pour la suite de l'analyse.

La séquence *EcoRIR2F4-1* a été séparée en deux parties à partir du site de restriction de *EcoRI*, générant les fragments *EcoRIR2F4-5'* (264 pb) et *EcoRIR2F4-3'* (947 pb). L'alignement (Fig. 34-A) puis l'assemblage (Fig. 34-B) de ces deux fragments avec le consensus de 380 pb *ConsAacutmar-sp1* a permis d'obtenir la première séquence reconstruite d'une longueur totale de 1268 pb, nommée *Re-1*. La comparaison par BlastX de *Re-1* avec les séquences dans banques de données montre qu'elle possède une homologie avec des transposases de type MLE, et notamment avec celles de la betterave *Beta vulgaris* (45 %) et de la diatomée marine *Pseudo-nitzschia multiseriis* (42 %). Cette première iPCR a donc bien permis d'identifier les régions bordant le fragment de 380pb *ConsAacutmar-sp1*, et d'allonger la séquence connue du gène de la transposase de MLE de plus de 880 pb.



**Figure 34 : A) Schéma de l'alignement de la séquence *ConsAacutmar-sp1* avec les fragments *EcoRIR2F4-5'* et *EcoRIR2F4-3'*. B) L'assemblage des trois fragments, dans l'ordre, *EcoRIR2F4-5'*, *ConsAacutmar-sp1* et *EcoRIR2F4-3'* a permis de reconstruire la séquence nommée *Re-I*.**

1	tttgtagacatatcaatgccctgaagccgacactcactcctgaacacatagtagcaagactggagtttaccatgagtaaagtagacca
1	F V R H I N A L K P T L T P E H I V A R L E F T M S K V D P
91	aatatggtaaggaatggaagtcggaagtacttaccaatgttagatgaggttcataattgatgagaagtggtccacatgaccaaggacaat
31	N M V R N G S P K Y L P M L D E V H <u>I</u> <u>D</u> <u>E</u> <u>K</u> <u>W</u> <u>F</u> H M T K D N
181	cagggtgttcattctgtcaaagaacgaggaaccacccaaaaggactgtgcagcacaagtggtacattggcaagtgatgttctgtgtgca
61	Q V F I L S K N E E P P K R T V Q H K C Y I G K V M F L C A
271	caagcaaggccaagaatggtagatggcgaatgtgggatggaagattggcatttggcatttggatattgttgagaagcacaagaaga
91	Q A R P R M V D G R M W D G K I G I W P F G Y V G E A Q R R
361	tcaaaacacagacccagaggagcaccggtttggaaaatgagagcggttaacaagaggtgtactgccgatacatgttgacaaggtcatt
121	S K H R P R G A P V W K N E S V K Q E V Y C R Y M L D K V I
451	cctgccattatagacaagttcccatgagatacttgaaagaggggtcaaaatccagcaagacaatgcaaaatcccatattcctgacaat
151	P A I I D K F P M R Y L K E G V K <u>I</u> <u>Q</u> <u>Q</u> <u>D</u> <u>N</u> <u>A</u> K S H I P D N
541	ggtgaagaatggttagatgccatgcaaagcatttacgatgagtagacaagctgaatattcagatatatactcaaccggctcaatcaccagac
181	G E E W L D A M Q S I Y D E Y K L N I Q I Y T Q P A Q <u>S</u> <u>P</u> <u>D</u>
631	acaaacatcaatgatgttggcattctttagaagcatccaaatgctttattatgaagcagcaccaacaacgacttgcctcatcaagaaa
211	<u>T</u> <u>N</u> <u>I</u> <u>N</u> <u>D</u> <u>L</u> <u>A</u> <u>F</u> <u>F</u> R S I Q M L Y Y E A A P T N D F A L I K K
721	ggtgagatgcataattggaagtatccagtttaacaagctcaatcagatgtattgacttaccaagctgtcttaacatgatcattgaacacg
241	V E D A Y W K Y P V N K L N Q M Y <u>Z</u> L T K A V L T <u>Z</u> S L N T
811	aggagacaacttttactcaattcctcacatgaacaaagaggcgttggaaggaccggccgaagtctctcagaggaggaataacttcatca
271	R E T T F T Q F L T <u>Z</u> T K R A W E G P A E V S Q R R N T S S
901	aaatggaacaaaagaacagacagaccaagataattgactattttgacattacgtctttccaattgtggttcttgaatcaggaccta
301	K W N K R T E Q T K I L T I L T L R L S N C G F L E S G T <u>Z</u>
991	gctactcggctagaccataattctcaggctacatcaaacacagctaactgcgactggtccatccgggttctttcgattttgcagtggtc
331	A T R L D H N S Q A T S N T A N C D W S I R V L S I L Q W L
1081	atttgtccaacaaaagggttcaggttaggaaaaaaagcaagctgacagcacaacttctggctcttggtgacttcttcttatcgccacact
361	I C P T K R F R L G K K S K L T A Q L L A L G L L S Y R H T
1171	ctttcattccattcgaagatacgccttggtgtgtgtgcctctcttctgttcttcttcttcttcttcttcttcttcttcttcttcttctt
391	L S F H S K D T P W C C V P L L C S F L F I P L I F L F P C

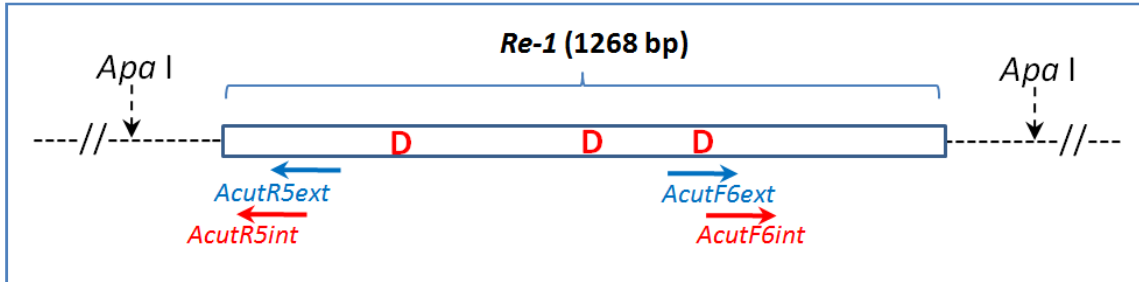
**Figure 35 : Traduction de la séquence *Re-1* en acides aminés.**

Les trois acides aspartiques sont indiqués en rouge et soulignés. La séquence de 380 pb de *ConsAacutmar-spl* est en rouge et les deux parties allongées en noir. Les motifs caractéristiques de la triade catalytique apparaissent en grisé. Z : codon stop.

Après traduction en protéines, la séquence *Re-1* a une longueur de 420 acides aminés (Fig. 35). Les motifs entourant les trois D de la triade catalytique ont pu être précisés pour les deux premiers et identifiés *de novo* pour le troisième. Ainsi le premier motif est "IDEKWF", le second "IQQDNA" et le troisième "SPDTNINDLAFF". La distance entre le deuxième et troisième D est de 43 acides aminés. Il y a 4 codons stop au niveau de l'extrémité 3' de *Re-1*. En se basant sur la structure générale du gène de la transposase de MLE, il manque dans la séquence *Re-1* plusieurs motifs protéiques typiques de la région 5' tels que les motifs Hélice-Tour-Hélice (HTH) ou encore le signal de localisation nucléaire (NLS), ce qui indique qu'il est nécessaire de poursuivre l'allongement de cette extrémité. Parallèlement, l'allongement de la région 3' a également été poursuivi. La séquence *Re-1* a donc servi de matrice pour choisir une enzyme et des amorces nécessaires à un nouvel allongement de la séquence par iPCR.

## 2.2. Allongement de la séquence *Re-1* en régions 5' et 3'

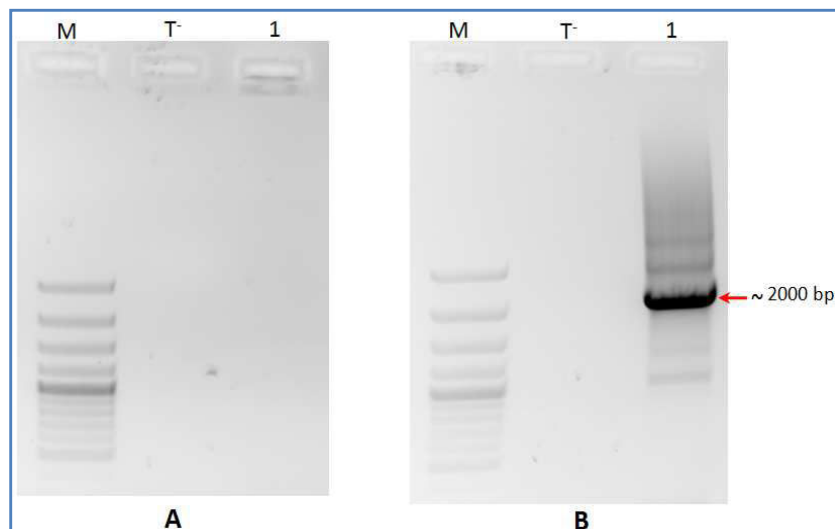
Pour poursuivre l'allongement de la séquence *Re-1* en régions 5' et 3', de l'ADN génomique d'*A. acutiuscula* a été digéré par l'enzyme de restriction *ApaI* (Fig. 36). Cette dernière ne coupe pas dans *Re-1* permettant ainsi d'obtenir en même temps les deux séquences flanquant *Re-1* par iPCR.



**Figure 36 : Schéma permettant de localiser les amorces ayant servi à l'allongement du fragment reconstruit *Re-1* de 1268 pb.**

L'enzyme *ApaI* qui ne coupe pas *Re-1* a été choisie. Les amorces *AcutR5ext* et *AcutF6ext* (pour externe, flèches bleues) ont été utilisées pour la iPCR n°1 et les amorces *AcutR5int* et *AcutF6int* (pour interne, flèches rouges) pour la iPCR n°2. La ligne pointillée représente les régions bordant la séquence *Re-1* en 5' et en 3'.

Après les deux amplifications successives, un fragment d'environ 2000 pb a été observé (Fig. 37-B).



**Figure 37 : Migration sur gel d'agarose des produits d'amplification pour l'obtention des fragments flanquant la séquence *Re-1*.**

**A** - Puits 1 : produit de la première iPCR utilisant comme matrice l'ADNg digéré par l'enzyme *ApaI* et le couple d'amorces externes *AcutR5ext* et *AcutF6ext*. **B** - Puits 1 : produit de la seconde iPCR utilisant comme matrice le produit de la iPCR précédente et le couple d'amorces internes *AcutR5int* et *AcutF6int*. M : marqueur de taille GeneRuler 100 pb. T<sup>-</sup> : témoin négatif (iPCR réalisée sans ADNg).

Ce dernier a été cloné, deux clones ont été séquencés et nommés *ApaIR5F6-1* et *ApaIR5F6-2*. Ils ont une taille respective de 1982 et 1983 pb et sont identiques à 99,3 %. Pour l'analyse, seul le clone *ApaIR5F6-1* a été utilisé. Au niveau du site de l'enzyme *ApaI*, il a été séparé en deux fragments *ApaR5F6-5'* et *ApaR5F6-3'* qui ont été alignés avec *Re-1* (Fig. 38-A).

En région 5', un bon alignement du fragment *ApaR5F6-5'* avec l'extrémité 5' de *Re-1* a permis d'allonger ce dernier d'environ 300 pb (Fig. 38-B). En région 3', le fragment *ApaIR5F6-3'* de 1579 pb et la séquence *Re-1* présentent un bon alignement sur une longueur de 220 bases suivi par un très mauvais alignement des 640 dernières bases de *Re-1*. A partir de ces résultats, deux séquences *Re-2.1* et *Re-2.2* ont pu être établies. La séquence *Re-2.1* a été assemblée en associant depuis la région 5' vers la région 3' : *ApaIR5F6-5'*, *Re-1* et *ApaIR5F6-3'*. La séquence *Re-2.2* n'a été allongée qu'en région 5' en associant uniquement les fragments *ApaIR5F6-5'* et *Re-1*. Ces deux séquences reconstruites partagent une partie identique de 1155 pb à l'extrémité 5', suivie d'une partie variable en région 3' (Fig. 39).

La comparaison par BlastX de *Re-2.1* et *-2.2* avec les banques de données a permis de confirmer qu'elles possèdent bien une homologie avec des transposases de type MLE, et notamment avec celle de la betterave *Beta vulgaris* (37 et 39 %, respectivement) et de la diatomée marine *Pseudo-nitzschia multiseriis* (39 %).



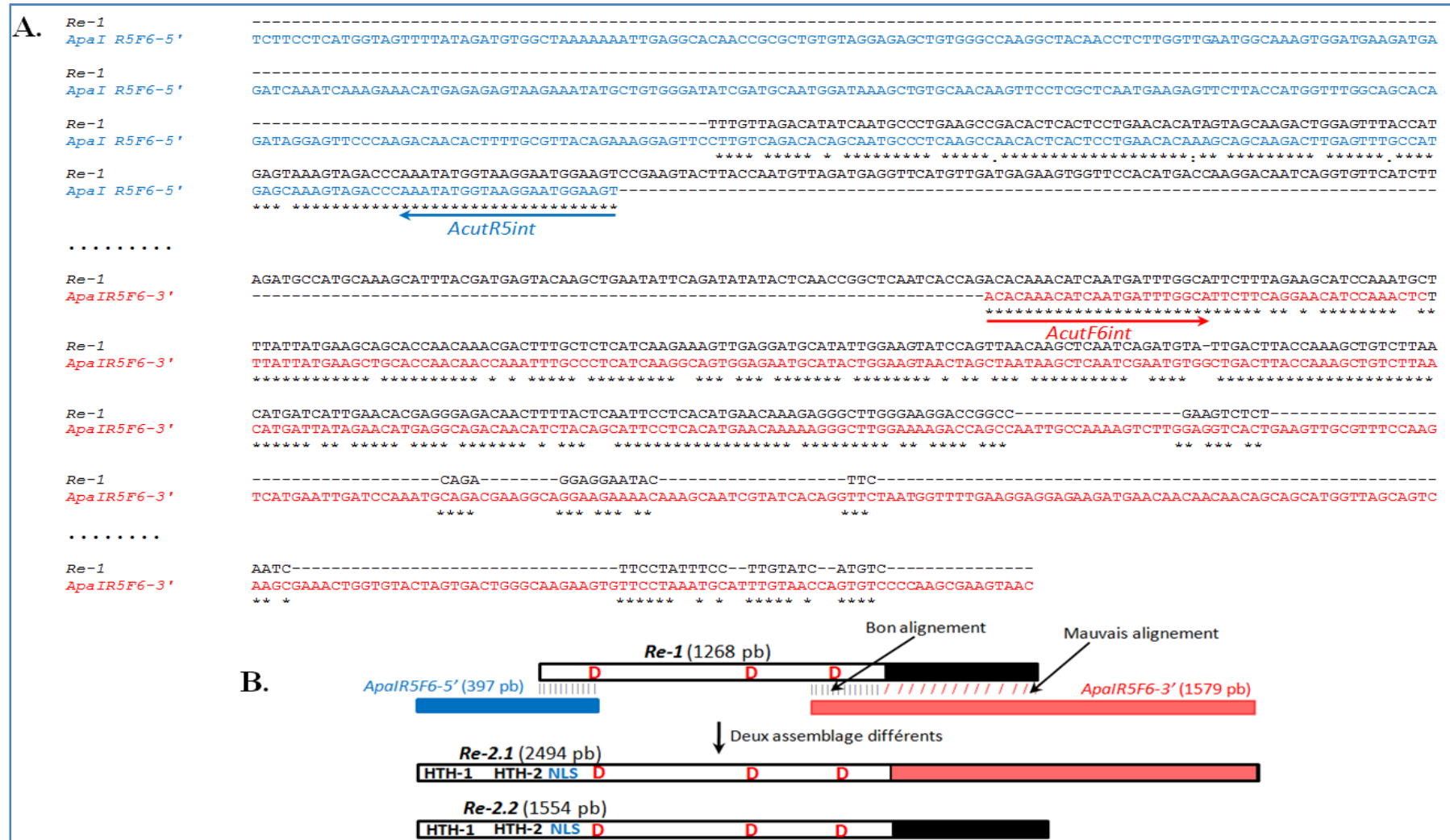


Figure 38 : A) Schéma de l'alignement de la séquence *Re-1* avec les fragments *ApaIR5F6-5'* et *ApaIR5F6-3'*. B) L'assemblage de *ApaIR5F6-5'* et *ApaIR5F6-3'* avec *Re-1*, d'une part et de *ApaIR5F6-5'* avec *Re-1* d'autre part a permis de reconstruire deux séquences différentes.

Re-2.1 : TCTTCCTCATGGTAGTTTATAGATGTGGCTAAAAAATTGAGGCACAACCGCGCTGTGTAGGAGAGCTGTGGGCC : 76  
 Re-2.2 : TCTTCCTCATGGTAGTTTATAGATGTGGCTAAAAAATTGAGGCACAACCGCGCTGTGTAGGAGAGCTGTGGGCC : 76

Re-2.1 : AAGGCTACAACCTCTTGGTTGAATGGCAAAGTGGATGAAGATGAGATCAAATCAAAGAAACATGAGAGAGTAAGAA : 152  
 Re-2.2 : AAGGCTACAACCTCTTGGTTGAATGGCAAAGTGGATGAAGATGAGATCAAATCAAAGAAACATGAGAGAGTAAGAA : 152

Re-2.1 : ATATGCTGTGGGATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAATGAAGAGTTCTTACCATGG : 228  
 Re-2.2 : ATATGCTGTGGGATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAATGAAGAGTTCTTACCATGG : 228

Re-2.1 : TTTGGCAGCACAGATAGGAGTTTCCAAGACAACACTTTTGGGATACAGAAAGGAGTTCCTTGTGACACACAGCAAT : 304  
 Re-2.2 : TTTGGCAGCACAGATAGGAGTTTCCAAGACAACACTTTTGGGATACAGAAAGGAGTTCCTTGTGACACACAGCAAT : 304

Re-2.1 : GCCCTCAAGCCAACACTCACTCCTGAACACAAAGCAGCAAGACTTGAATTTGCCATGAGCAAAGTAGACCCAAATA : 380  
 Re-2.2 : GCCCTCAAGCCAACACTCACTCCTGAACACAAAGCAGCAAGACTTGAATTTGCCATGAGCAAAGTAGACCCAAATA : 380

Re-2.1 : TGGTAAGGAATGGAAGTCCGAAGTACTTACCAATGTTAGATGAGGTTTCATGTTGATGAGAAGTGGTTCCACATGAC : 456  
 Re-2.2 : TGGTAAGGAATGGAAGTCCGAAGTACTTACCAATGTTAGATGAGGTTTCATGTTGATGAGAAGTGGTTCCACATGAC : 456

Re-2.1 : CAAGGACAATCAGGTGTTTCATCTTGTCAAAGAACGAGGAACCAACCAAAAGGACTGTGCAGCACAAAGTGTACATT : 532  
 Re-2.2 : CAAGGACAATCAGGTGTTTCATCTTGTCAAAGAACGAGGAACCAACCAAAAGGACTGTGCAGCACAAAGTGTACATT : 532

Re-2.1 : GGCAAAGTGATGTTCTTGTGTGCACAAGCAAGGCCAAGAATGGTAGATGGCCGAATGTGGGATGGAAGATTTGGCA : 608  
 Re-2.2 : GGCAAAGTGATGTTCTTGTGTGCACAAGCAAGGCCAAGAATGGTAGATGGCCGAATGTGGGATGGAAGATTTGGCA : 608

Re-2.1 : TTTGGCCATTTGGATATGTTGGAGAAGCACAAGAAGATCAAACACAGACCCAGAGGAGCACCAGTTTGGAAAAA : 684  
 Re-2.2 : TTTGGCCATTTGGATATGTTGGAGAAGCACAAGAAGATCAAACACAGACCCAGAGGAGCACCAGTTTGGAAAAA : 684

Re-2.1 : TGAGAGCGTTAAACAAGAGGTGTACTGCCGATACATGTTGGACAAGGTTCATTCCTGCCATTATAGACAAGTTCCCC : 760  
 Re-2.2 : TGAGAGCGTTAAACAAGAGGTGTACTGCCGATACATGTTGGACAAGGTTCATTCCTGCCATTATAGACAAGTTCCCC : 760

Re-2.1 : ATGAGATACTTGAAAGAGGGGGTCAAAATCCAGCAAGACAATGCAAAAATCCCATATTCTTGACAATGGTGAAGAAT : 836  
 Re-2.2 : ATGAGATACTTGAAAGAGGGGGTCAAAATCCAGCAAGACAATGCAAAAATCCCATATTCTTGACAATGGTGAAGAAT : 836

Re-2.1 : GGTTAGATGCCATGCAAAGCATTTACGATGAGTACAAGCTGAATATTCAGATATATACTCAACCGGCTCAATCACC : 912  
 Re-2.2 : GGTTAGATGCCATGCAAAGCATTTACGATGAGTACAAGCTGAATATTCAGATATATACTCAACCGGCTCAATCACC : 912

Re-2.1 : AGACACAAACATCAATGATTTGGCATTCTTACGAAACATCCAAACTCTTTATTATGAAGCTGCACCAACAACCAAA : 988  
 Re-2.2 : AGACACAAACATCAATGATTTGGCATTCTTACGAAACATCCAAACTCTTTATTATGAAGCTGCACCAACAACCAAA : 988

Re-2.1 : TTTGCCCTCATCAAGGCAGTGGAGGATGCATCTGGAAGTAACTAGCTAATAAGCCCAATCGAATGTGGCTGACTT : 1064  
 Re-2.2 : TTTGCTCTCATCAAGAAAGTTGAGGATGCATATGGAAGTATCCAGTTAAACAAGCTCAATCAGATGTA-TTGACTT : 1063

Re-2.1 : ACCAAAGCTGTCTTAACATGATATTGAACATGAGGCAGACAACCTCTACAGCATTCCTCACATGAACAAAAGGG : 1140  
 Re-2.2 : ACCAAAGCTGTCTTAACATGATATTGAACATGAGGCAGACAACCTTTACTCTATTCTCCTCACATGAACAAAAGGG : 1139

Re-2.1 : CTTGCAAAAGACCGCCAAATTGCCAAAAGTCTTGGAGGTCACTGAAGTTGCGTTTCCAAGTCATGAATTGATCCAA : 1216  
 Re-2.2 : CTTGGCAAGACCGCC-----GAGTCTCT----- : 1165

```

      1220      *      1240      *      1260      *      1280      *
Re-2.1 : ATGCAGACGAAGGCAAGAAAGAAACAAAGCAATCGTATCACAGCTTCATATGGTTTGAAGGAGGAGAAGATGAAC : 1292
Re-2.2 : ---CAGA-----GGAGGAATAC-----TTC----- : 1182

      1300      *      1320      *      1340      *      1360
Re-2.1 : AACACAACAGCAGCATGGTTAGCAGTCATCAAAATACAATGACTGACAAGATGACCAAAACACGATGACCAATCA : 1368
Re-2.2 : -----ATCAAA-----ATC-----GACCAAA----- : 1198

      *      1380      *      1400      *      1420      *      1440
Re-2.1 : AGGCCAATACCAGGCCAATCCCCACACAAACACACCGATAGCACCACAAAGAGATACAGAACAAAGAACTTACACC : 1444
Re-2.2 : -----AGAACACA-----ACAGACCAAGA----- : 1217

      *      1460      *      1480      *      1500      *      1520
Re-2.1 : GTATTAGTAAATAGAGAGCAAATAACAAAGGAATCCCATGGTCTTTCATCTAAAACAGAGAGCTGCTGCTGTTGG : 1520
Re-2.2 : -TATGA-----CTA-----TTTGA----- : 1232

      *      1540      *      1560      *      1580      *
Re-2.1 : AATCAAGTTTGTGATTGAAAATGAGCTCTGCAGCGTCTGCTGCTGAGGAGTTTAGTGTTGGGGAAAACCTTTCGCTCC : 1596
Re-2.2 : CATACGTCCT----- : 1243

      1600      *      1620      *      1640      *      1660      *
Re-2.1 : GTGGCGCGAAGCGAACTGAAAGGATTCACTGATTGTTAAAAATGCTCCACGGTGTCTTTTGACACAGCTTCT : 1672
Re-2.2 : -----TCCAATTC-----TG-----GTTTCCTTGATCAGC----- : 1269

      1680      *      1700      *      1720      *      1740
Re-2.1 : TATGAATGATCCAGCTAACCCAGCTGCTGTAGAATTTGTACCAGTTCTACGTGGGTCTTGTTTCGGTGCAGGCAG : 1748
Re-2.2 : -----GACCTAAGCTA-CTCGCTA----- : 1288

      *      1760      *      1780      *      1800      *      1820
Re-2.1 : GACTGATGATAATTCCTGGACAGAGGGGGTAGGCGTGTAACACATGCGCGAGGACGAGGATGGACCCATCTCGT : 1824
Re-2.2 : GACC-----ATAATTC-----CAG-----GCTACAT-----CAACACACCTAAC : 1324

      *      1840      *      1860      *      1880      *      1900
Re-2.1 : TTTGATGCTGGCAAGAAAGAAAAAGCGCCTCTTTTGGTGGTTTCATCCGTTGGGCTATGTTGCTACGGCATAAACAT : 1900
Re-2.2 : TGGGAC-----TGGTCCATCC-----GGGTCTCTT-----CGAT : 1353

      *      1920      *      1940      *      1960      *
Re-2.1 : CATACCAAGCACACCCCCAAAAATACAATCGACATCAGTATATGATACATCCATCACCCCTGCTTCGAACCTTCAATA : 1976
Re-2.2 : TTTC-----TTCAT-----AGTG-----GCTCAT-----TTC-----TCCAAACA : 1378

      1980      *      2000      *      2020      *      2040      *
Re-2.1 : TGACAGGTCTCTGGTCAACACACAATAAAAAAATTGATCACCGCATAGCGCCGCGCGCCCAACGTTTAAAGCAG : 2052
Re-2.2 : -AAAGGTTCA--GGTAG-----GAAAAAAA--AGC-----AGC-----AAGCTC : 1411

      2060      *      2080      *      2100      *      2120
Re-2.1 : ACGATGCTGCGCCATTCATTATGATTCAACAACCTCTGTAATTATCATGTCGTATAGGCGCGGATATCTTCGATTTC : 2128
Re-2.2 : ACAGC-----ACAACCTCTG-----GC-----TCTTGGAGTTCT : 1440

      *      2140      *      2160      *      2180      *      2200
Re-2.1 : ATTTGGGTGGGATGTGCGGATAGTTTCATCATGTGACTAGTCGGAGAGGGATAGAGCCAAATCTCTATAAAATTT : 2204
Re-2.2 : TTCTT-----ATC-----GCCACA-CTCT----- : 1458

      *      2220      *      2240      *      2260      *      2280
Re-2.1 : GGTTCAGAAAAATTCATAAATTATGATTACAGGGAAGAAAGCGATGAACGTGCGACGAGTGCCACGACAACAT : 2280
Re-2.2 : --TTCA-----TTCAT-----TCGA----- : 1472

      *      2300      *      2320      *      2340      *
Re-2.1 : GGGACACGGGGGCAATGCTGCTGGAATACATGCACTCCCTAAACCCGACGGCCCTCTCAACACAACAATGGCACTT : 2356
Re-2.2 : AAGATACGCG-----TTCCTGTTCT-----TCT-----GCCTCTC----- : 1502

      2360      *      2380      *      2400      *      2420      *
Re-2.1 : TGTCAATAATCTTCTGGGTTAGGAGCAGACGTGGCACAAGGCCTGGGGACTTCTCTAAAGGAACTGGTGTAC : 2432
Re-2.2 : TGTCT-----TCTTTCTCTCTCA-----TCCCTTAAATC----- : 1531

      2440      *      2460      *      2480      *
Re-2.1 : TAGTGACTGGGCAAGAAGTGTTCCTAATGCATTTGTAAACAGTGTCCTCCCAAGCGAAGTAAC : 2494
Re-2.2 : -----TTCCT--ATTTCCTGTATC--ATGTC----- : 1554

```

**Figure 39 : Alignement des séquences reconstruites *Re-2.1* et *Re-2.2*.**  
Elles sont identiques sur 1155 pb en région 5', la région 3' est variable.

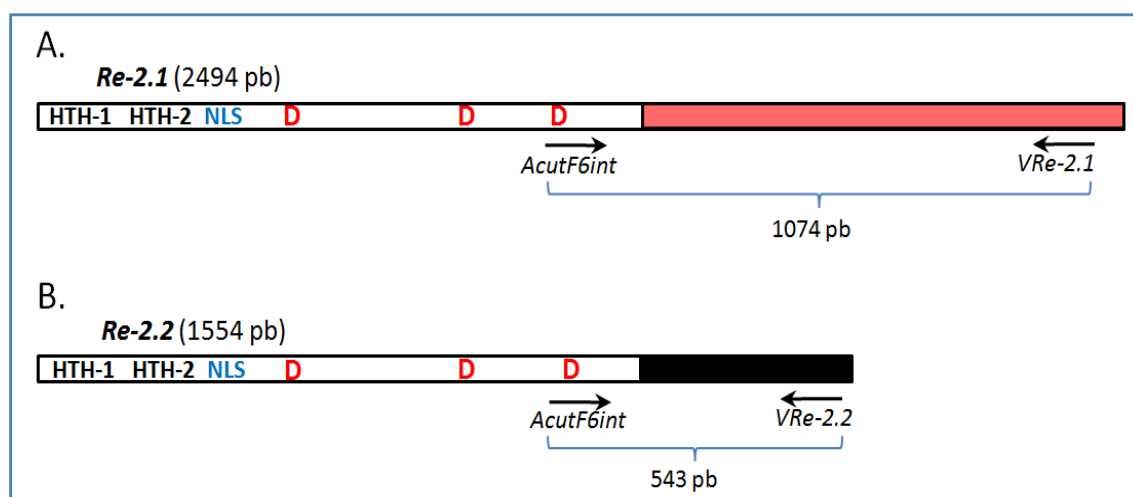
Dans la région 5' de *Re-2.1* et -2.2, l'analyse a permis d'identifier un motif HTH similaire à celui d'autres transposases disponibles dans les banques de données. En complément, la recherche du motif HTH a été effectuée en utilisant un logiciel spécifique "HTH prediction", le score obtenu étant faible (2,48- probabilité non significative) montre que ce motif est mal défini. Cependant, le logiciel de prédiction indique la présence d'un autre HTH en amont de celui détecté précédemment avec un score significatif de 3,25 (probabilité 50 %). Un signal de localisation nucléaire a été détecté, après les motifs HTH, grâce à l'utilisation du logiciel NLS Mapper. Ainsi en région 5', les séquences *Re-2.1* et -2.2 présentent, comme beaucoup de transposases, deux HTH et un NLS (Fig. 40). Toutefois la méthionine de départ n'a pas été trouvée, indiquant qu'il est nécessaire de poursuivre la marche le long de la séquence en direction de l'extrémité 5'.

1	cttcctcatggtagttttatagatgtggctaaaaaattgaggcacaaccgctgtgttaggagagctgtggccaaggctacaacctct
1	L P H G S F I D V A K K I E A Q P R C V G E L W A K A T T S
	<b>HTH-1</b>
91	tggttgaatggcaaagtggatgaagatgagatcaaatcaaagaacatgagagagtaagaaatatgctgtgggatatcgatgcaatggat
31	W L N G K V D E D E I K S K K H E R V R N M L W D I D A M D
181	gaagctgtgcaacaagttcctcgctcaatgaagagttcttaccatggtttggcagcacagataggagtttccaagacaacacttttgca
61	E A V Q Q V P R S M K S S Y H G L A A Q I G V S K T T L L R
	<b>HTH-2</b>
361	tacagaaaggagttccttgtcagacacagcaatgccctcaagccaacactcactcctgaacacaaagcagcaagacttgaatttgccatg
91	Y R K E F L V R H S N A L K P T L T P E H K A A R L E F A M
	<b>Bipartite NLS</b>
451	agcaaagtagacccaatatggttaaggatggaagtccgaagtacttaccatgttagatgagggttcattgttgatgagaagtgggtccac
121	S K V D P N M V R N G S P K Y L P M L D E V H V D E K W F H

**Figure 40 : Traduction en protéine de l'extrémité 5' identique des deux séquences *Re-2.1* et -2.2 et positionnement des motifs caractéristiques des transposases.**

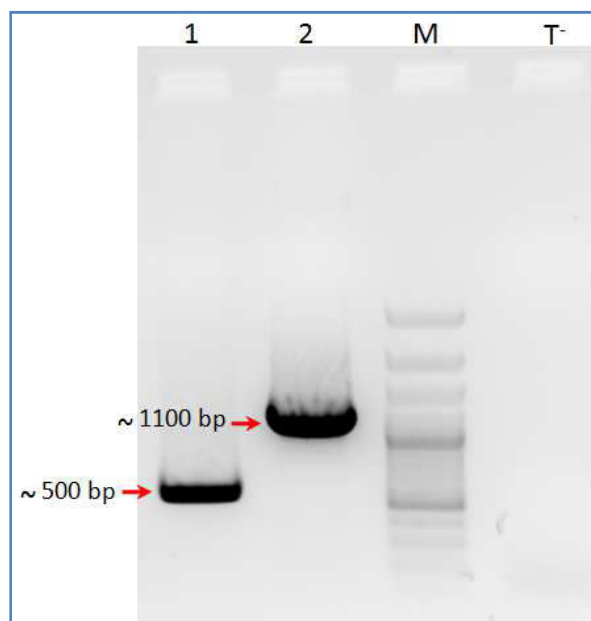
En région 5' des séquences *Re-2.1* et -2.2, deux motifs HTH (Hélice-Tour-Hélice) et un motif NLS (Signal de localisation nucléaire) ont été prédits par les logiciels HTH prédiction et NLS Mapper.

Le mauvais alignement d'une partie de l'extrémité 3' du fragment *ApaR5F6-3'* avec la séquence *Re-1* nous a conduit à vérifier l'existence des séquences reconstruites *Re-2* dans le génome d'*A. acutiuscula*. Pour cela, une amorce spécifique s'hybridant près de l'extrémité 3' de chacune des deux séquences *Re-2.1* et *Re-2.2* a été définie, il s'agit des amorces antisens *VRe-2.1* et *VRe-2.2* (Fig. 41). Pour un meilleur accrochage, les amorces antisens ont été définies quelques bases en amont de la fin de la séquence. L'amorce sens *AcutF6int* s'hybride, quant à elle, dans une zone très conservée du gène de la transposase sur le motif contenant le troisième acide aspartique de la triade catalytique. Chaque couple d'amorces doit permettre l'amplification du fragment allant de la partie terminale du gène de transposase jusqu'à l'extrémité 3' du fragment reconstruit et donc de vérifier l'existence de ce dernier dans le génome-hôte (Fig. 42). Dans les deux cas, les produits de PCR présentaient la taille attendue qui est de 1085 pb pour l'élément *Re-2.1* et 543 pb pour l'élément *Re-2.2*. Ces fragments ont été clonés puis séquencés. Leur amplification a permis de confirmer par alignement (Fig. 43 et 44) qu'il existe bien dans le génome d'*A. acutiuscula* deux éléments distincts, le premier ayant en extrémité 3' la séquence présente dans *Re-2.1* et le second la séquence 3' définie dans l'élément *Re-2.2*.



**Figure 41 : Schéma de la localisation des amorces utilisées pour vérifier la présence des séquences des éléments reconstruits *Re-2.1* et *Re-2.2*.**

Utilisation du couple d'amorces *AcutF6int* et *VRe-2.1* pour la mise en évidence de *Re-2.1* et du couple d'amorces *AcutF6int* et *VRe-2.2* pour celle de *Re-2.2*.



**Figure 42 : Amplifications à partir de l'ADNg d'*A. acutiuscula* des fragments de vérification des régions 3' des séquences *Re-2.1* et *-2.2*.**

Puits 1 : produit de PCR obtenu avec le couple d'amorces *AcutR6int* et *VRe-2.1* ; Puits 2 : produit de PCR obtenu avec le couple d'amorces *AcutR6int* et *VRe-2.2*. M : marqueur de taille Generuler 100 pb. T<sup>-</sup> : témoin négatif (PCR réalisée sans matrice ADNg).

Les séquences *Re-2.1* et *Re-2.2* ont une longueur de 2494 et 1554 pb, respectivement. Elles possèdent 1074 et 543 pb au-delà du 3<sup>ème</sup> D de la triade catalytique, il est donc possible qu'elles contiennent la fin de l'élément transposable, l'UTR3' et l'ITR 3'. En région 5', le codon start n'ayant pas encore été détecté, cette région doit encore être allongée pour obtenir la région antérieure de la transposase ainsi que l'UTR et l'ITR 5'. Une fois l'allongement effectué en 5', les deux régions seront comparées pour déceler la présence éventuelle des ITR. Si tel n'est pas le cas, la marche vers les deux extrémités sera alors reprise. Dans le paragraphe suivant la priorité est donnée à l'allongement de la région 5' des séquences *Re-2* (2.1 et 2.2).



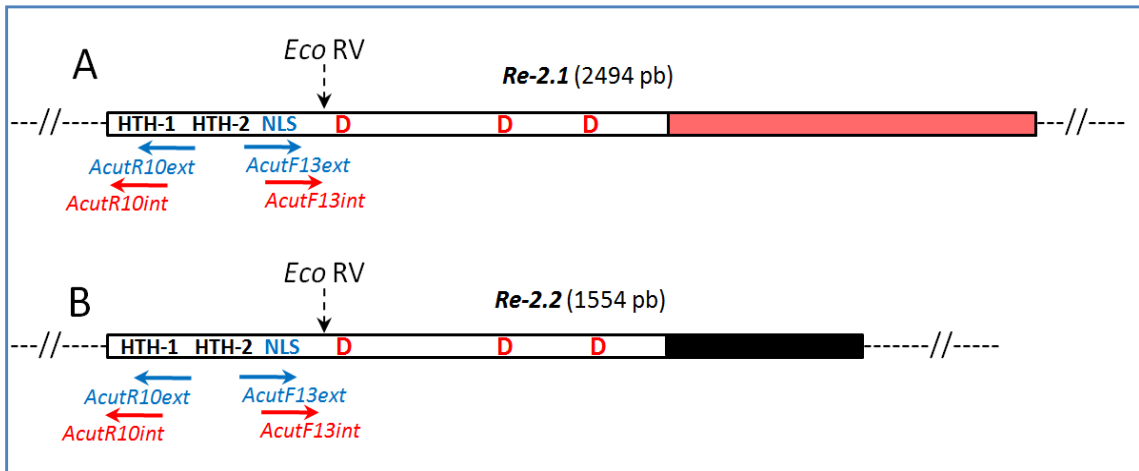
	<i>AcutF6int</i> →		
<i>F6int/V-Re2.1</i>	ACACAAACATCAATGATATGGCATTCTTCAGGAACATCCAAACTCTTTATTATGAAGCTGCACCAACAACCAATTTGCCCTCATCAAGGCAGTGGAGGATGCATACTGGAAGTAACTAG	120	
<i>Re-2.1</i>	ACACAAACATCAATGATTTGGCATTCTTCAGGAACATCCAAACTCTTTATTACGAGCTGCACCAACAACCAATTTGCCCTCATCAAGGCAGTGGAGGATGCATACTGGAAGTAACTAG	120	
	*****		
<i>F6int/V-Re2.1</i>	CTAATAAGCTCAATCGAATGTGGCTGACTTACCAAAGCTGTCTTAACATGATTATTGAACATGAGGCAGACAACATCTACAGCATTCCTTCACATGAACAAAAAGGGCTTGAAAAGACCA	240	
<i>Re-2.1</i>	CTAATAAGCCCAATCGAATGTGGCTGACTTACCAAAGCTGTCTTAACATGATTATTGAACATGAGGCAGACAACATCTACAGCATTCCTTCACATGAACAAAAAGGGCTTGAAAAGACCA	240	
	*****		
<i>F6int/V-Re2.1</i>	GCCAATTGCCAAAAGTCTTGAGGTCCTGAAGTTGCGTTTCCAAGTCATGAATTGATCCAAATGCAGACGAAAGGCAGGAAGAAAACAAAGCAATCGTATCACAGGTTCTAATGGTTTTG	360	
<i>Re-2.1</i>	GCCAATTGCCAAAAGTCTTGAGGTCCTGAAGTTGCGTTTCCAAGTCATGAATTGATCCAAATGCAGACGAAAGGCAGGAAGAAAACAAAGCAATCGTATCACAGGTTCTAATGGTTTTG	360	
	*****		
<i>F6int/V-Re2.1</i>	AAGGAGGAGAAGATGAACAACAACAACAGCAGCATGGTTAGCAGTCAATCAAATACAATGACTGACAAGATGACCAAAACACGATGACCAAAATCAAGGCCAATACCAGGCCAATCCCCACA	480	
<i>Re-2.1</i>	AAGGAGGAGAAGATGAACAACAACAACAGCAGCATGGTTAGCAGTCAATCAAATACAATGACTGACAAGATGACCAAAACACGATGACCAAAATCAAGGCCAATACCAGGCCAATCCCCACA	480	
	*****		
<i>F6int/V-Re2.1</i>	CAAAACACACCGATAGCACCACAAAGAGATACAGAAACAAGAACTTACACCGTATTTAGTTAAATGAGAGCAAAATAACAAGGAATCCCATGGTCTTTCATCTAAACAGAGAGCTGCTGC	600	
<i>Re-2.1</i>	CAAAACACACCGATAGCACCACAAAGAGATACAGAAACAAGAACTTACACCGTATTTAGTTAAATGAGAGCAAAATAACAAGGAATCCCATGGTCTTTCATCTAAACAGAGAGCTGCTGC	600	
	*****		
<i>F6int/V-Re2.1</i>	TGCTGGAATCAAGTTTTGATTGAAAATGAGCTCTGCAGCGTGCTGCTGCTGAGGAGTTTAGTGTGGGAAAAACTTTCGCTCCGTGGCGCGAAGCGAACTGAAAGGATTCTCTGATTGGT	720	
<i>Re-2.1</i>	TGTTGGAATCAAGTTTTGATTGAAAATGAGCTCTGCAGCGTGCTGCTGCTGAGGAGTTTAGTGTGGGAAAAACTTTCGCTCCGTGGCGCGAAGCGAACTGAAAGGATTCTCTGATTGGT	720	
	** *****		
<i>F6int/V-Re2.1</i>	TAAAAATGTCTCCACGGTGTTCTTTGCACCACGTTCTTATGAATGATCCACGCTAACCAGCTGCTGTAGAATTTGTACAGTTCTACGTGGGTCTGTTTCGGTGCAGGCAGGACTGA	840	
<i>Re-2.1</i>	TAAAAATGTCTCCACGGTGTTCTTTGCACCACGTTCTTATGAATGATCCACGCTAACCAGCTGCTGTAGAATTTGTACAGTTCTACGTGGGTCTGTTTCGGTGCAGGCAGGACTGA	840	
	*****		
<i>F6int/V-Re2.1</i>	TGAATAATTCTGGACCAGAGGGGGTAGGCGTGTAGCGACATGCGCGAGGACGAGGATGGACCCATCTCGTTTTGATGCTGGCAAGAAGAAAAAAGCGCCTCTTTTGGTGTTGATCCGT	960	
<i>Re-2.1</i>	TGAATAATTCTGGACCAGAGGGGGTAGGCGTGTAGCGACATGCGCGAGGACGAGGATGGACCCATCTCGTTTTGATGCTGGCAAGAAGAAAAAAGCGCCTCTTTTGGTGTTGATCCGT	960	
	*****		
<i>F6int/V-Re2.1</i>	TGGGCTATGTGCTACGGCATAACAATCATACCAAGCACACCCCAAAATACAATCGACATCAGTATATGATACACTCATCACCTTGCTTCGAACCTTCAATATGACAGGTCTCT-----	1074	
<i>Re-2.1</i>	TGGGCTATGTGCTACGGCATAACAATCATACCAAGCACACCCCAAAATACAATCGACATCAGTATATGATACACTCATCACCTTGCTTCGAACCTTCAATATGACAGGTCTCTCGGTCA	1080	
	*****		
			← <i>V-Re2.1</i>

**Figure 43 : Alignement du fragment *F6int/VRe-2.1* avec la région 3' du fragment reconstruit *Re-2.1*.**  
 Sur la partie amplifiée, les deux fragments ont une identité de 99,53 %.

	<i>AcutF6int</i>	
<i>Re-2.2</i>	ACACAAACATCAATGATTTGGCATTCTTTAGAAGCATCCAAATGCTTTATTATGAAGCAGCACCAACAAACGACTTTGCTCTCATCAAGAAAGTTGAGGATGCATATTGGAAGTATCCAG	120
<i>F6int/V-Re2.2</i>	ACACAAACATCAATGATTTGGCATTCTTTAGAAGCATCCAAATGCTTTATTATGAAGCAGCACCAACAAACGACTTTGCTCTCATCAAGAAAGTTGAGGATGCATATTGGAAGTATCCAG	120
	*****	
<i>Re-2.2</i>	TTAACAAGCTCAATCAGATGTATTGACTTACCAAAGCTGTCTTAACATGATCATTGAACACGAGGGAGACAACTTTTACTCAATTCTCACATGAACAAAGAGGGCTTGGGAAGGACCGG	240
<i>F6int/V-Re2.2</i>	TTAACAAGCTCAATCAGATGTATTGACTTACCAAAGCTGTCTTAACATGATCATTGAACACGAGGGAGACAACTTTTACTCAATTCTCACATGAACAAAGAGGGCTTGGAAAAGACCGG	240
	*****	*****
<i>Re-2.2</i>	CCGAAGTCTCTCAGAGGAGGAATACTTCATCAAAATGGAACAAAAGAACAGAACAGACCAAGATATTGACTATTTTGACATTACGCTTTTCCAATTGTGGTTTCCTTGAATCAGGGACCT	360
<i>F6int/V-Re2.2</i>	CCGAAGTCTCTCAGAGGAGGAATACTTCATCAAAATGGAACAAAAGAACAGAACAGACCAAGATATTGACTATTTTGACATTATGCTTTTCCAATTGTGGTTTCCTTGAATCAGGGACCT	360
	*****	*****
<i>Re-2.2</i>	AAGCTACTCGGCTAGACCATAATTCTCAGGCTACATCAAACACAGCTAACTGCGACTGGTCCATCCGGGTTCTTTCGATTTTGAGTGGCTCATTTGTCCAACAAAAGGTTTCAGGTTAG	480
<i>F6int/V-Re2.2</i>	AAGCTACTCGGCTAGACCATAATTCTCAGGCTACATCAAACACAGCTAACTGCGACTGGTCCATCCGGGTTCTTTCGATTTTGAGTGGCTCATTTGTCCAACAAAAGGTTTCAGGTTAG	480
	*****	
<i>Re-2.2</i>	GAAAAAAAGCAAGCTGACAGCACAACTTCTGGCTCTTGGACTTCTTTCTTATCGCCACACTCTTTCATTCCATTGAAAGATACGCCTTGGTGTGTGCCTCTCTTGTGTTCTTTTC	600
<i>F6int/V-Re2.2</i>	GAAAAAAAGCAAGCTGACAGCACAACTTCTGGCTCTTGGACTTCTTTCTTATCGCCACACTC-----	543
	*****	
	<i>V-Re2.2</i>	

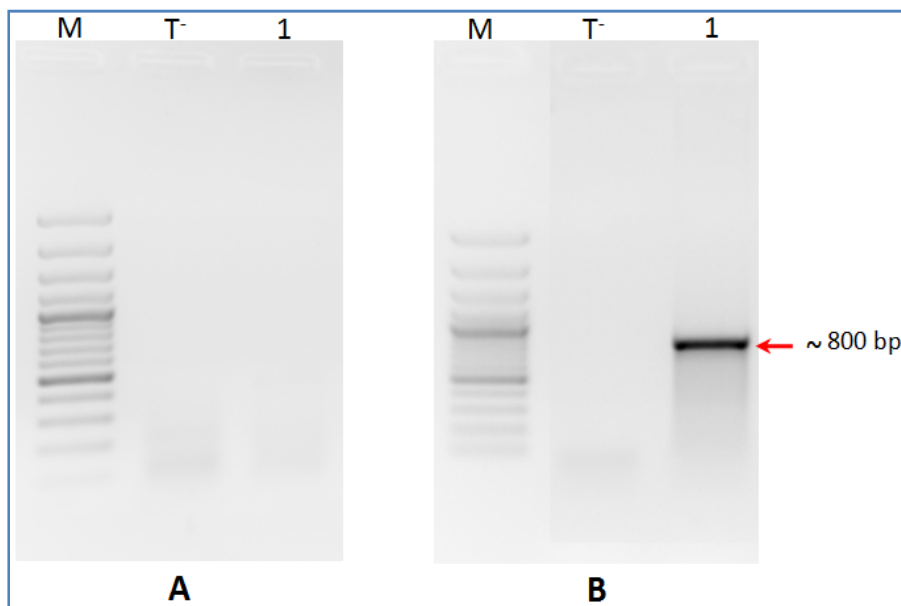
**Figure 44 : Alignement du fragment de vérification *F6int/VRe-2.2* avec la région 3' de la séquence reconstruite *Re-2.2*.**  
 Sur la partie amplifiée, les deux fragments possèdent une identité de 99,45 %.





**Figure 45 : Schéma permettant de localiser les amorces ayant servi à l'allongement de l'extrémité 5' de *Re-2.1* (A) et *Re-2.2* (B).**

L'enzyme *EcoRV* qui coupe les séquences *Re-2* au niveau du domaine de liaison à l'ADN a été choisie. Les flèches en bleu et en rouge représentent les deux couples d'amorces nichées utilisés. Les amorces *AcutR10ext* et *AcutF13ext* (pour externe) ont été utilisées pour l'iPCR n°1 et les amorces *AcutR10int* et *AcutF13int* (pour interne) pour l'iPCR n°2. La lignepointilléereprésenteles régions bordant les séquences *Re-2*.



**Figure 46 : Amplification par iPCR du fragment bordant l'extrémité 5' des séquences *Re-2*.**

**A-** Puits 1 : amplification par iPCR en utilisant l'ADN génomique digéré par l'enzyme *EcoRV*, comme matrice et le couple d'amorces externes *AcutR10ext* et *AcutF13ext*. **B -** Puits 1 : amplification par iPCR en utilisant comme matrice le produit de l'iPCR précédente et le couple d'amorces internes *AcutR10int* et *AcutF13int*. M : marqueur de taille 100 pb. T<sup>-</sup> : témoin négatif (iPCR réalisée sans matrice).

### **2.3. Allongement des séquences *Re-2* vers l'extrémité 5'**

Afin de poursuivre l'allongement des séquences *Re-2* vers la région 5', nous avons choisi l'enzyme *EcoRV* qui coupe à l'intérieur de *Re-2*, et défini deux nouveaux couples d'amorces nichées, *AcutR10ext* et *AcutF13ext*, et *AcutR10int* et *AcutF13int*, s'hybridant dans la partie 5' des séquences *Re-2* (Fig. 45).

Après l'iPCR, un fragment d'environ 800 pb a été obtenu et cloné, et deux clones ont été séquencés (Fig. 46). Ces derniers ont été nommés *EcoRVR10F13-1* et *EcoRVR10F13-2*, ils sont identiques à 98,6 %, de ce fait, seul le clone *EcoRVR10F13-1* a été utilisé pour l'analyse. Il a été séparé en deux parties, *EcoRVR10F13-5'* et *EcoRVR10F13-3'*, au niveau du site de restriction de l'enzyme *EcoRV*, puis le fragment *EcoRVR10F13-5'* a été aligné avec la région 5' consensus des deux séquences *Re-2.1* et *2.2* (Fig. 47-A).

L'alignement montre que le fragment *EcoRVR10F13-5'* a permis l'allongement des séquences *Re-2.1* et *Re-2.2* de plus de 600 pb vers la région 5' (Fig. 47-A). L'assemblage de *EcoRVR10F13-5'* avec *Re-2.1* et *Re-2.2* respectivement, a permis de reconstruire deux nouvelles séquences, nommées *Re-3.1* et *Re-3.2* (Fig. 47-B). La longueur totale de ces dernières est respectivement de 3117 pb et de 2177 pb. Dans la séquence *Re-3.1*, nous pouvons observer le gène de transposase complet avec un codon start en position 455 et des parties de 454 pb et 1182 pb bordant ce gène en région 5' et 3', respectivement. C'est dans ces dernières que les ITR et le TSD, qui est un dinucléotide TA, ont été recherchés par une méthode développée par JD Rouault (cf Matériel et Méthodes). Cette recherche a permis de détecter la présence des ITR uniquement dans la séquence *Re-3.1* (Fig. 48-B). Après leur détection, les ITR de *Re-3.1* ont été alignés avec la séquence *Re-3.2* afin de comprendre pourquoi ils n'ont pu être localisés dans celle-ci. En fait, cette dernière ne contient pas l'ITR-3' d'où l'impossibilité de détecter les répétitions inversées constituant les ITR par les méthodes bioinformatiques. Dans la séquence *Re-3.1*, une région de 27 pb en orientation inversée bordée par un dinucléotide TA a été mise en évidence à chaque extrémité. Les ITR-5' et -3' ont donc une taille de 27 pb. Ils ne sont pas identiques car ils montrent quatre différences en positions 2, 15, 21 et 22 (Fig. 47).

ITR-5'	:	CTCCCTCTGGTCCAAATTAACATCA	:	27
ITR-3'	:	CCCCCTCTGGTCCAGAATTATTCATCA	:	27
Consensus	:	CYCCCTCTGGTCCARAATTAWWCATCA	:	

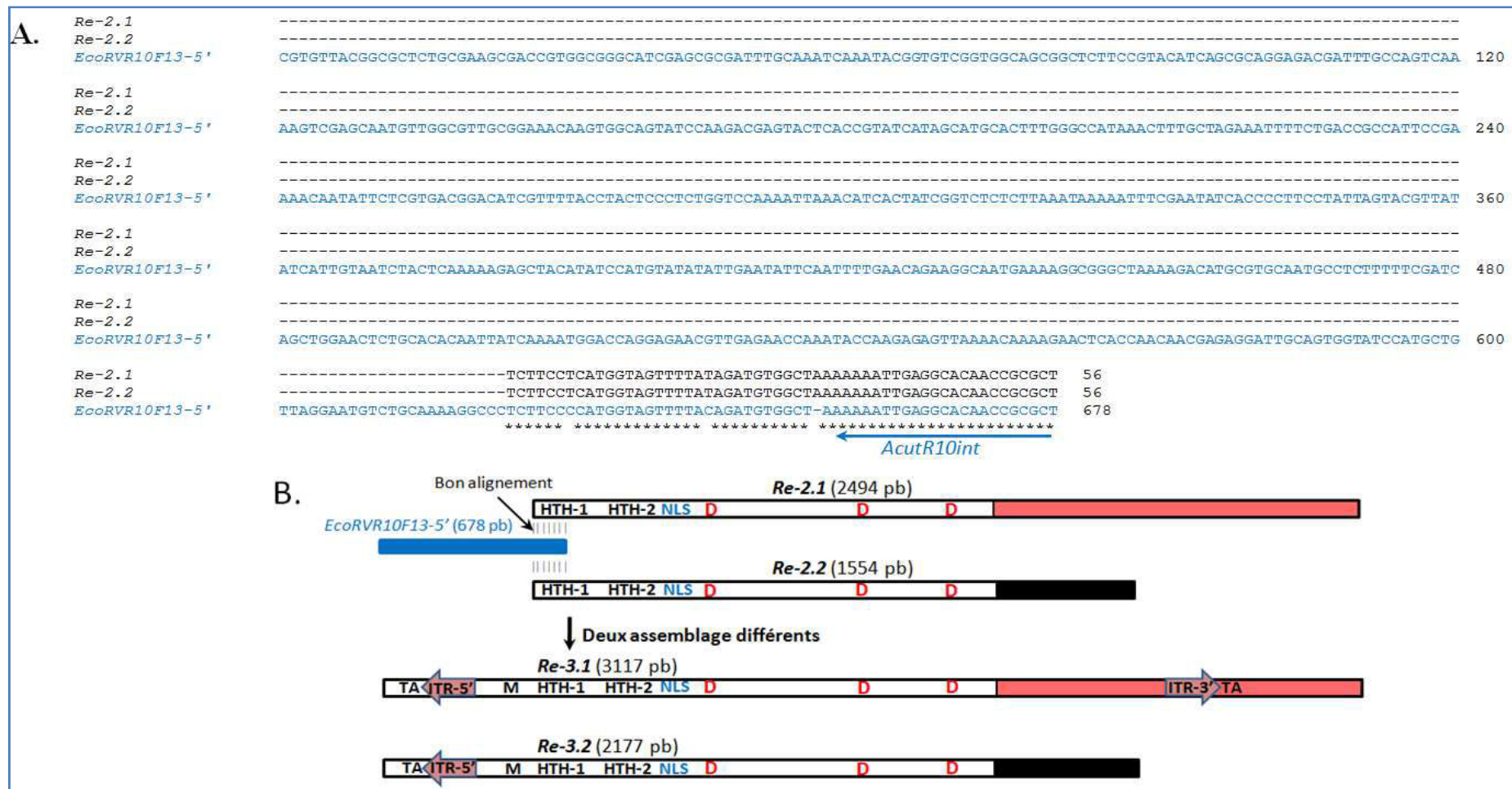
**Figure 47 : Comparaison des séquences des ITR-5' et -3' détectées dans l'élément reconstruit *Re-3.1* d'*A. acutiuscula*.**

Quatre différences sont présentes en positions : 2, 15, 21 et 22.

L'existence des ITR et la duplication du site-cible ont permis de délimiter le transposon pleine taille. Celui-ci a été nommé *Aacutmar1-re* (pour *Amphora acutiusculamariner* n°1 reconstruit) et possède une longueur totale de 2130pb.

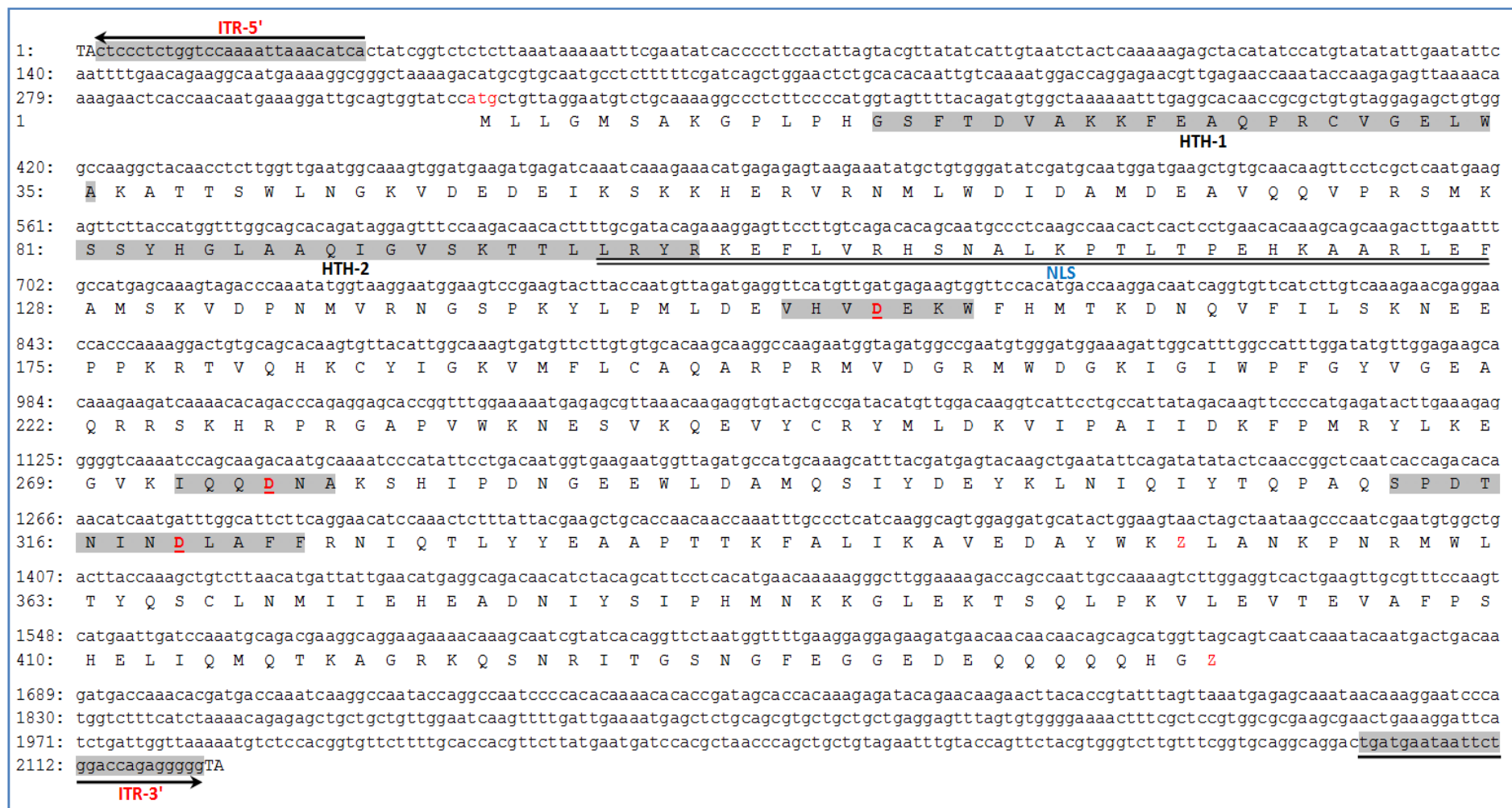
Les étapes successives ayant conduit à l'obtention de la séquence complète *Aacutmar1-re* sont récapitulées dans la figure 49. Celle-ci présente de façon schématique les séquences reconstruites dérivant les unes des autres et marquant l'avancée le long de la séquence recherchée pour aboutir à un transposon pleine taille.

La traduction in silico d'*Aacutimar1-re* a montré un cadre de lecture qui, après analyse par BlastX, donne un gène de transposase sur toute sa longueur soit 1341 pb. La traduction conceptuelle de l'ORF a révélé une transposase putative de type *mariner* de 447 acides aminés avec un codon start en position 320 et un codon stop en position 1661.



**Figure 48 : A) Alignement du fragment *EcoRVR10F13-5'* avec les deux séquences *Re-2.1* et *Re-2.2*. B) L'assemblage du fragment *EcoRVR10F13-5'* avec *Re-2.1* et *-2.2* a permis de former les séquences *Re-3.1* et *Re-3.2*, respectivement.**

Dans ces séquences sont détectés depuis la région 5' vers 3' : les motifs HTH 1 et 2 (en noir), le NLS (en bleu) et les trois D de la triade catalytique (en rouge). En plus des motifs présents dans les séquences *Re-2*, les ITR-5' et -3' (flèches rouges) pour *Re-3.1* et uniquement l'ITR 5' pour *Re-3.2* sont indiqués.



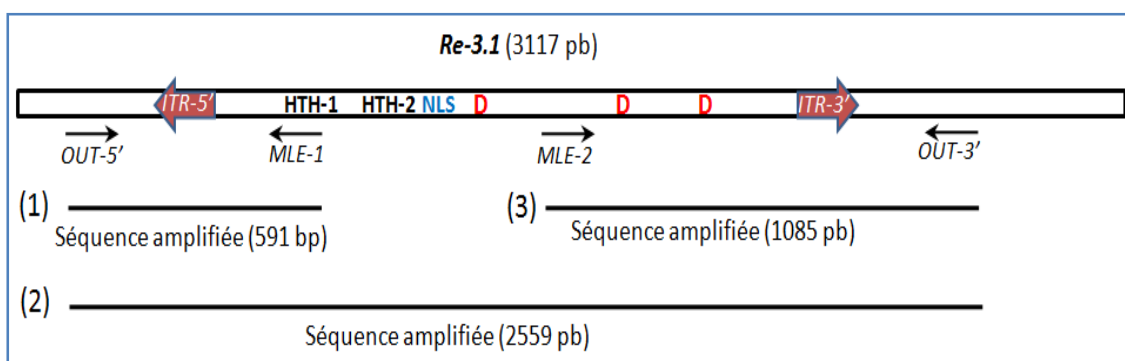
**Figure 49 : Séquences nucléique et protéique du transposon *Aacutmar1*-recomplet identifié à partir de la séquence reconstruite *Re-3.1*.**

Les motifs de la transposase : les deux motifs HTH sont grisés et notés, le motif NLS est souligné par deux traits, les trois motifs caractéristiques de la triade catalytique sont grisés et chaque acide aspartique est en rouge et souligné. Les ITR sont grisés et soulignés aux deux extrémités.

### 3. Caractérisation du transposon *Aacutmar1* dans le génome d'*A. acutiuscula*

#### 3.1. Caractérisation des ITR d'*Aacutmar1*

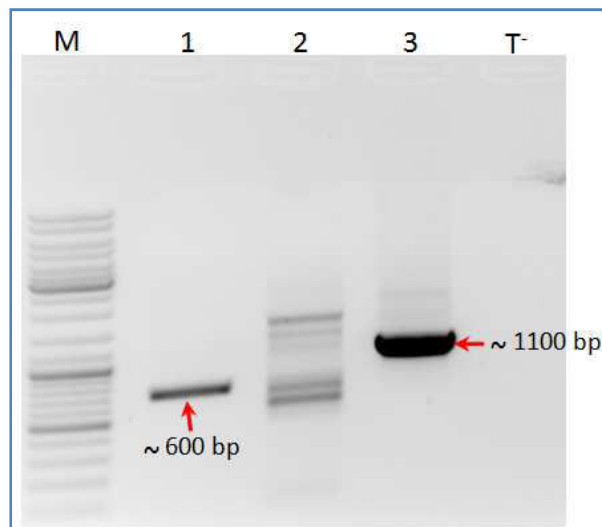
Afin de confirmer la présence de l'élément complet (*Aacutmar1-re*) dans le génome d'*A. acutiuscula*, des amorces ont été définies à partir des ITR. Toutefois, avant d'établir ces amorces, il semblait important de vérifier la séquence des ITR dans le génome hôte. Pour cela, des amorces s'hybridant de part et d'autre des régions contenant les ITR-5' et -3' ont été définies. Ainsi, deux amorces positionnées au-delà des ITR (*OUT-5'* et *OUT-3'*) et deux amorces situées dans la séquence de l'ORF (*MLE-1* et *MLE-2*) ont été définies (Fig. 50). Le couple *OUT-5'* et *MLE-1* a permis d'amplifier la zone contenant l'ITR-5', le couple *MLE-2* et *OUT-3'*, celle contenant l'ITR-3'.



**Figure 50 : Schéma de l'élément *Re-3.1* montrant la localisation des amorces définies pour vérifier les ITR de l'élément *Aacutmar1* dans le génome d'*A. acutiuscula*.**

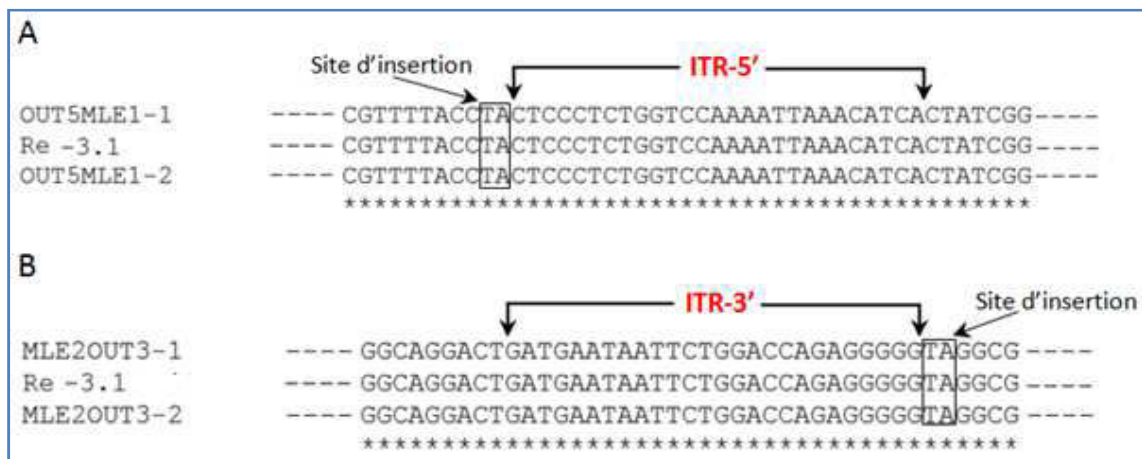
Les flèches sous la séquence indiquent la position des couples d'amorces. Le couple d'amorces *OUT-5'* et *MLE-1* a été défini pour amplifier la région contenant l'ITR-5', le couple d'amorces *MLE-2* et *OUT-3'* pour amplifier la région contenant l'ITR-3'. Les traits indiquent la taille attendue de chaque séquence amplifiée.

Après amplifications séparées des régions contenant les ITR-5' et -3' (Fig. 51), les produits de PCR présentent une bande unique de forte intensité et ayant la taille attendue, environ 600 pb pour la séquence contenant l'ITR-5' et environ 1100 pb pour celle contenant l'ITR-3' ont été clonés. Pour chaque amplification, deux clones ont été séquencés et analysés, ils ont été nommés en région 5' *OUT5'MLE1-1* et *1-2* et en région 3' *MLE2OUT3'1-1* et *1-2*. Ces séquences ont été alignées avec les ITR détectés dans l'élément reconstruit *Re-3.1* confirmant ainsi les séquences des ITR-5' et -3'. Ils ont une longueur de 27 pb et sont bordés par le TSD (Fig. 52).



**Figure 51 : Migration sur gel d'agarose des produits de PCR contenant les ITR de l'élément *Aacutmar1* dans le génome d'*A. acutiuscula*.**

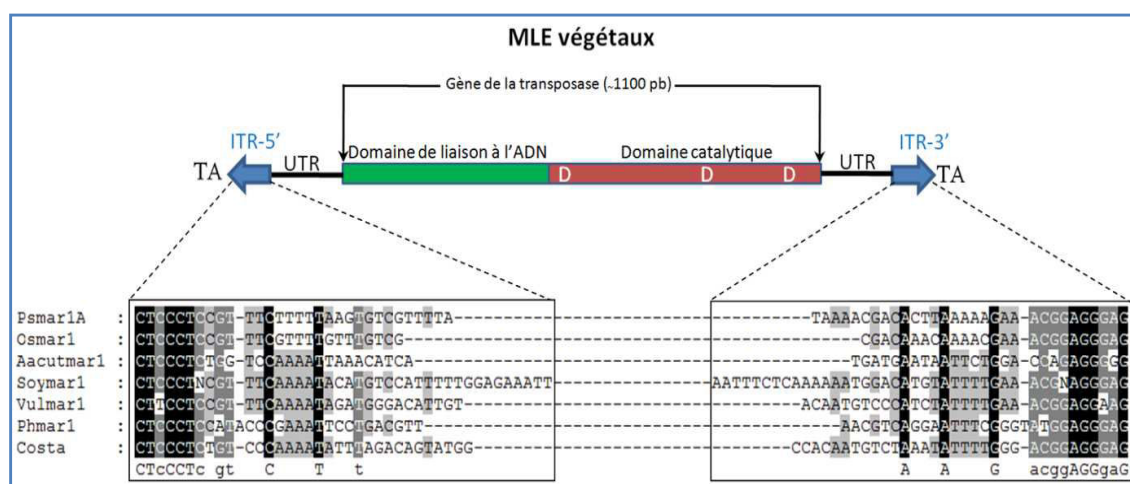
Puits 1: produit de PCR obtenu avec les amorces *OUT-5'*/*MLE-1* contenant l'ITR-5'. Puits 2 : produits de PCR obtenu par les amorces *OUT-5'*/*OUT-3'*. Puits 3 : produit de PCR obtenu avec les amorces *MLE-2*/*OUT-3'* contenant l'ITR-3'. M : marqueur de taille Generuler Mix. T : témoin négatif (PCR réalisée sans ADNg).



**Figure 52 : Alignement des séquences des ITR-5' (A) et -3' (B) obtenues par iPCR dans l'élément reconstruit *Re-3.1* avec ceux obtenus lors de la vérification par PCR à partir de l'ADNg d'*A. acutiuscula*.**



La comparaison des ITR d'*Aacutmar1-re* avec ceux d'éléments complets de végétaux décrits dans la littérature montre des homologues avec *Soymar1* du soja (Jarvick et Lark, 1998), *Costa* de la diatomée *Pseudo-nitzschia multiseries* (numéro d'accès JX013907, séquence non publiée), *Vulmar1* de la betterave (Jacobs et al., 2004), *Osmar1* du riz (Tarchini et al., 2000), *Phmar1* du bambou (Tang et al., 2010) et *Psmar1* du petit pois (Macas et al., 2005). Cependant, l'identité nucléotidique est limitée en région 5' aux 18 premiers nucléotides des ITR avec un consensus 5'-CTCCCTCCGT(C/T)CAAAT-3' (Fig. 53) et en région 3' aux 23 premières paires de bases avec un consensus 5'-ATTG(G/A)ACGGAGGGAG-3'.



**Figure 53 : Comparaison des ITR-5' et -3' d'*A. acutiuscula* avec ceux des éléments complets disponibles dans la littérature et dans les banques de données.**

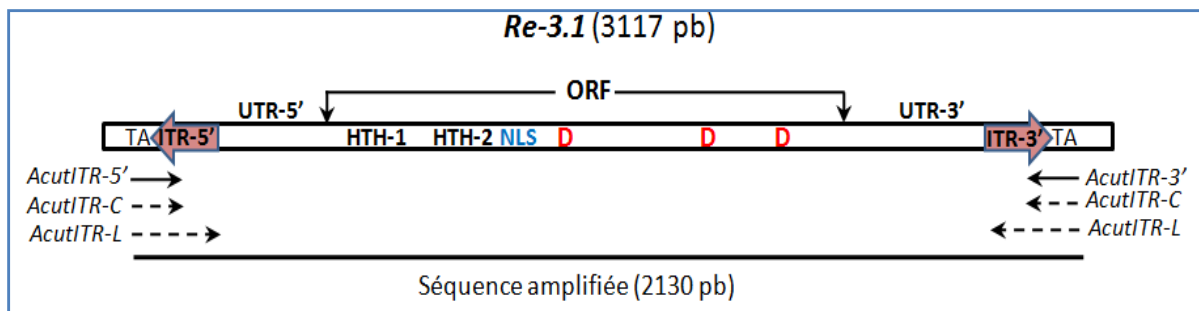
## **3.2. Caractérisation de séquences complètes des transposons *Aacutmar1***

### **3.2.1. Identification des MLE *Aacutmar1* complets à l'aide des amorces définies sur les ITR**

L'objectif ici était d'amplifier des séquences complètes des transposons *Aacutmar1*. A cette fin, des amorces ont été définies sur les ITR (Fig. 54). Les séquences des ITR caractérisées dans le paragraphe précédent sont identiques à celles de la séquence *Re-3.1*. A l'extrémité de chaque ITR, un dinucléotide TA, caractéristique de l'insertion d'un MLE (TSD), a été trouvé également. Les séquences des TSD et des ITR ont servi à établir des amorces destinées à amplifier des éléments *Aacutmar1* complets. Afin d'optimiser les chances d'obtenir des amplifications de qualité, plusieurs amorces ont été

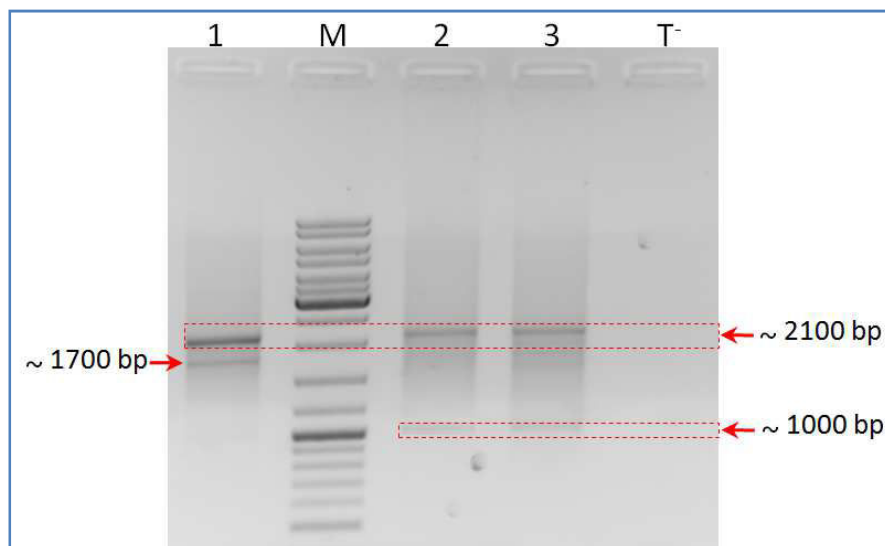


définies : un couple d'amorces de 20 pb spécifiques des ITR-5' et -3' et nommées *AcutITR-5'* et *AcutITR-3'*, deux amorces dégénérées susceptibles de se fixer chacune sur les deux ITR, appelées respectivement *AcutITR-C* (C pour Courte) et *AcutITR-L* (L pour Longue), elles ont des longueurs respectives de 20 et 29 pb. Les amorces *AcutITR-L* de 29 pb couvrent toute la séquence de l'ITR et incluent celle du TSD. Une même amorce pouvant se fixer sur chaque ITR et permettre l'amplification d'un élément complet, les amorces dégénérées *AcutITR-C* et *AcutITR-L* ont été utilisées seules.



**Figure 54 : Schéma de l'amplification des éléments complets *Aacutmar1* avec les amorces spécifiques et dégénérées définies sur les ITR.**

Ces amorces s'hybrident soit sur les 20 soit sur les 29 premiers nucléotides des ITR. En se basant sur l'élément reconstruit *Aacutmar1-re*, les amorces *AcutITR* doivent permettre l'amplification d'un produit de PCR d'une longueur d'environ 2130 pb. Les amorces *AcutITR5'* et *AcutITR3'* figurées en trait plein sont spécifiques, *AcutITR-C* et *AcutITR-L* figurées en pointillés sont dégénérées.

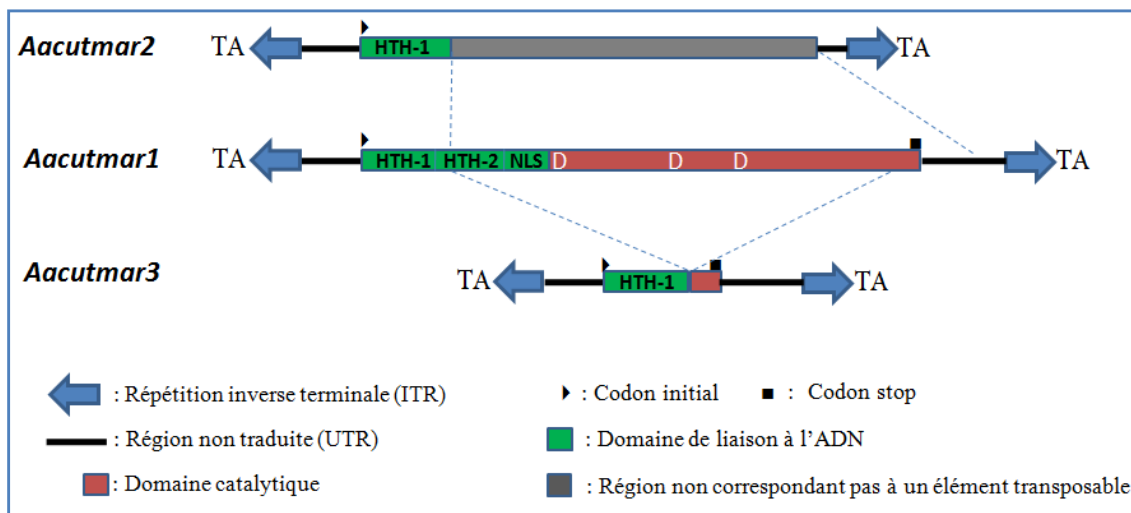


**Figure 55 : Migration sur gel d'agarose des produits de PCR obtenus avec les amorces établies sur les ITR d'*Aacutmar1-re*.**

Puits 1: produits de PCR obtenus avec les amorces spécifiques *AcutITR-5'*/*AcutITR-3'*. Puits 2 : produits de PCR obtenus par l'amorce dégénérée *AcutITR-C*. Puits 3 : produits de PCR obtenus

avec l'amorce dégénérée *AcutITR-L*. M : marqueur de taille Generuler Mix. T<sup>-</sup> : témoin négatif (PCR réalisée sans ADNg).

La PCR utilisant le couple d'amorces spécifiques *AcutITR-5'* et *AcutITR-3'* a donné deux bandes de 2100 et 1700 pb (puits 1, Fig. 55). Les amplifications réalisées d'une part avec l'amorce dégénérée courte, *AcutITR-C* (puits 2, Fig. 55) et d'autre part avec l'amorce longue, *AcutITR-L* (puits 3, Fig. 55) ont toutes deux donné deux bandes d'environ 2100 et 1000 pb. Les différentes amorces ont donc permis d'amplifier trois types de bandes : celles à la taille attendue (environ 2100 pb) et celles, plus courtes, d'une taille de 1700 et 1000 pb pouvant représenter des éléments *Aacutmar1* tronqués. Les 3 types de bandes ont été clonés. Deux clones de chaque fragment (2100, 1700 et 1000 bp) ont été séquencés. Après analyse, ont été obtenues des séquences pleine taille, nommées *Aacutmar1.1* et *1.2* de 2114 pb, des séquences plus courtes, *Aacutmar2.1* et *2.2* de 1753 pb et *Aacutmar3.1* et *3.2* de 1006 pb. Les analyses par BlastX ont montré que les séquences *Aacutmar1* correspondent à des MLE mais pas les séquences *Aacutmar2* et *3*, toutefois les 3 types de séquences ont été comparés entre elles en utilisant *Aacutmar1.1* et *Aacutmar1.2* comme références. En fait, les séquences *Aacutmar2* et *Aacutmar3* présentent une délétion d'une partie importante du gène de la transposase, de 1615 pb et 1112 pb, respectivement. C'est pour cette raison que leur analyse par BlastX n'a pas permis la détection de transposons de type MLE. Dans les séquences *Aacutmar2*, la partie délétée a été remplacée par un fragment d'ADN qui ne correspond à aucune séquence répertoriée dans les banques de données. Il a également été recherché la correspondance possible de cette dernière avec un élément transposable. Le résultat s'étant avéré négatif, cette partie ne correspond pas à un ET inséré dans un élément tronqué d'*Aacutmar1*.



**Figure 56 : Structure et organisation des différents éléments *Aacutmar* présents dans le génome d'*A. acutiuscula*.**

### **3.2.2. Établissement d'un consensus pour les éléments *Aacutmar1***

Les séquences *Aacutmar1* sont des transposons de type *mariner* pleine taille. Les transposons étant généralement répétés dans le génome hôte, il nous a semblé important d'identifier plusieurs séquences afin de pouvoir établir un consensus. Il a été décidé de porter de deux à douze le nombre d'éléments *Aacutmar1* séquencés. Pour ce faire, quatre réactions de PCR indépendantes ont été menées dans le but de diminuer la probabilité de ré-amplifier une seule et même séquence. Pour chaque PCR, 3 clones ont été envoyés au séquençage. Un total de 12 séquences nucléotidiques d'*Aacutmar1* a ainsi été obtenu et analysé en détail. Chaque séquence a été nommée *Aacutmar1.X*, X étant compris entre 1 et 12. Leur taille varie de 2094 à 2132 pb. Dans tous les clones, le codon d'initiation ATG et le codon stop TAG ont été identifiés. Parmi les 12 clones obtenus, six contiennent un cadre de lecture ouvert codant une transposase putative (Tab. 8), les six autres possédant des mutations ponctuelles, des insertions et/ou des délétions générant des codons stop et/ou des décalages du cadre de lecture. Sur leur longueur totale, les 12 clones partagent une identité nucléotidique de 90 % à 99 % et leurs ORF une identité de 91 % à 99 % en nucléotides et de 83 à 99 % en acides aminés (Tab. 9). Ces résultats montrent un polymorphisme assez important des éléments *Aacutmar1*. Afin de compléter l'analyse, le contenu en GC de l'ensemble des séquences *Aacutmar1* a été calculé ; il est situé entre 42 et 45 % pour les éléments pleine taille et pour l'ORF. Le taux de GC en position 3 des codons de l'ORF est proche de 40 %, ce qui montre un usage préférentiel des codons utilisant davantage l'adénine et la thymine que la guanine et la cytosine.

**Tableau 8 : Tableau récapitulant les caractéristiques des 12 séquences complètes des éléments *Aacutmar1* et de leur consensus obtenus chez la diatomée *A. acutiuscula*.**

Pour chaque séquence sont précisés : son nom, sa longueur, la présence ou non de codons stop, de délétions de décalages du cadre de lecture, le pourcentage de GC sur l'élément pleine taille, sur l'ORF et en position 3 des codons de l'ORF.

Nom	Longueur de MLE (pb)	Longueur de l'ORF pb (acides aminés)	ORF ininterrompue	Décalages du cadre de lecture	% GC		
					MLE	ORF	3ème position
<i>Aacutmar1.1</i>	2114	1473 (491)	1		44,75	45,62	41,87
<i>Aacutmar1.2</i>	2114	1473 (491)	1		44,84	45,69	42,07
<i>Aacutmar1.3</i>	2114	1473 (491)	1		44,89	45,69	41,87
<i>Aacutmar1.4</i>	2130	1473 (491)	0	1 codon stop interne	44,74	45,49	42,89
<i>Aacutmar1.5</i>	2132	1482 (494)	0	insertion de 5 pb, 3 codons stop internes	42,45	42,44	37,58
<i>Aacutmar1.6</i>	2116	1476 (492)	0	3 codons stop internes	41,73	42,82	38,34
<i>Aacutmar1.7</i>	2111	1470 (490)	1		44,95	45,58	42,16
<i>Aacutmar1.8</i>	2094	1449 (483)	0	2 délétions de 24 et 15 pb, 1 codon stop	42,74	43,34	38,64
<i>Aacutmar1.9</i>	2114	1473 (491)	1		44,75	45,82	42,07
<i>Aacutmar1.10</i>	2113	1473 (492)	0	3 codons stop internes, insertion de 6 pb	41,84	42,91	39,63
<i>Aacutmar1.11</i>	2114	1474 (492)	0	3 codons stop internes, insertion de 6 pb	42,05	42,3	43,9
<i>Aacutmar1.12</i>	2130	1473 (491)	1		44,93	45,69	42,89
<i>ConsAacutmar1</i>	2114	1491 (491)	1		44,23	45,13	41,68

**Tableau 9 : Comparaison deux à deux des 12 séquences complètes d'*Aacutmar1*.**

Pourcentage d'homologie sur la séquence complète et, entre parenthèses, sur l'ORF, en nucléotides et en acides aminés.

	<i>Aacutmar1.1</i>	<i>Aacutmar1.2</i>	<i>Aacutmar1.3</i>	<i>Aacutmar1.4</i>	<i>Aacutmar1.5</i>	<i>Aacutmar1.6</i>	<i>Aacutmar1.7</i>	<i>Aacutmar1.8</i>	<i>Aacutmar1.9</i>	<i>Aacutmar1.10</i>	<i>Aacutmar1.11</i>
<i>Aacutmar1.2</i>	99 (99-99)										
<i>Aacutmar1.3</i>	99 (99-99)	99 (99-99)									
<i>Aacutmar1.4</i>	94 (94-94)	94 (94-93)	93 (94-94)								
<i>Aacutmar1.5</i>	91 (92-87)	91 (92-87)	91 (92-87)	91 (91-86)							
<i>Aacutmar1.6</i>	91 (92-89)	91 (92-89)	91 (92-89)	91 (92-88)	90 (91-84)						
<i>Aacutmar1.7</i>	99 (99-99)	99 (99-98)	99 (99-99)	94 (94-92)	91 (92-87)	91 (91-88)					
<i>Aacutmar1.8</i>	91 (92-89)	91 (92-89)	91 (92-89)	91 (91-88)	95 (95-88)	90 (91-88)	91 (92-88)				
<i>Aacutmar1.9</i>	99 (99-99)	99 (99-99)	99 (99-99)	94 (94-93)	91 (92-87)	91 (92-89)	99 (99-99)	91 (92-89)			
<i>Aacutmar1.10</i>	91 (91-89)	91 (91-89)	91 (91-89)	91 (92-87)	90 (91-84)	99 (99-99)	91 (92-87)	90 (91-86)	91 (91-88)		
<i>Aacutmar1.11</i>	91 (91-90)	91 (91-89)	91 (92-90)	91 (92-80)	90 (91-84)	99 (99-99)	91 (92-88)	90 (91-87)	91 (91-89)	99 (99-99)	
<i>Aacutmar1.12</i>	94 (94-94)	93 (94-93)	94 (94-94)	99 (99-99)	90 (91-86)	91 (92-88)	93 (94-92)	90 (91-88)	94 (94-93)	91 (91-82)	91 (92-91)

### **3.2.3. Caractérisation des séquences complètes *Aacutmar1***

L'alignement des 12 séquences d'*Aacutmar* a permis de déduire un consensus nommé *ConsAacutmar1*. Ce dernier a une longueur totale de 2114 pb. Comme mentionné plus haut, les ITR ont une longueur de 27 pb. Les UTR ont une longueur moyenne de 150 pb en région 5' et 430 pb en région 3'. Dans ces séquences, la recherche d'une boîte TATA avec le logiciel HCTata a révélé une boîte putative en position 124 à 129 pb (Fig. 57)

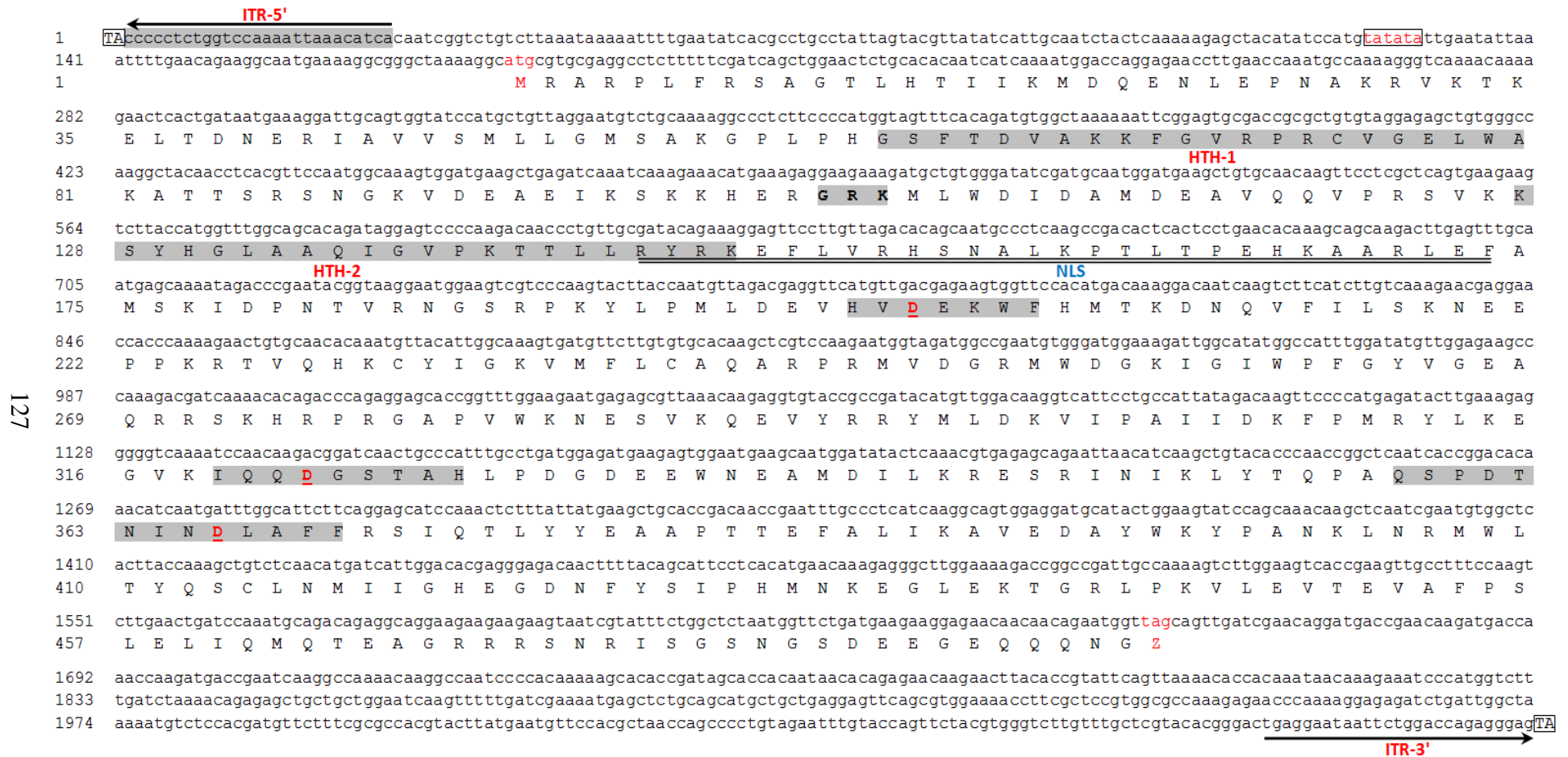
Après traduction en protéines de *ConsAacutmar1*, il a été observé un gène à cadre de lecture ouvert correspondant à une transposase de 491 acides aminés. Celle-ci a été appelée *AaTnp* (pour *Amphora acutiuscula* Transposase), elle présente un codon d'initiation de traduction (ATG) en position 180 et un codon stop en position 1653.

#### ***Recherche de motifs Hélice-Tour-Hélice (HTH)***

La séquence consensus *AaTnp* a été placée dans le logiciel de prédiction de motifs Hélice-Tour-Hélice. La structure prédite a révélé la présence d'un HTH, "GSFTDVAKKFGVRPRCVGELWA", des acides aminés 60 à 81 dans le domaine de liaison à l'ADN avec un score de 3,25. Ce score indique que la protéine contient un HTH avec une probabilité proche de 50 %. Un second motif HTH, "KSYHGLAAQIGVSKTTLLRYR", positionné du 128<sup>ème</sup> au 148<sup>ème</sup> acide aminé, est prédit mais avec un score faible de 2,02, ce qui ne permet pas d'affirmer la présence de ce second motif.

#### ***Recherche de signaux de localisation nucléaire (NLS)***

La recherche de signaux de localisation nucléaire a été effectuée dans toute la séquence *AaTnp* à l'aide du logiciel NLS Mapper. Une structure NLS, "RYRKEFLVRHSNALKPTLTPEHKAARLEF" est prédite avec un score relativement élevé de 3,2 indiquant la présence probable de ce type de signal. Ce motif chevauche sur 4 acides aminés le deuxième HTH (Fig. 57). Le motif NLS est localisé dans le domaine de liaison à l'ADN de la protéine du 145<sup>ème</sup> au 173<sup>ème</sup> acide aminé. (Numéro d'acide aminé dans la transposase).



**Figure 57 : Séquence consensus du transposon *ConsAacutmar1*.**

La traduction in silico en protéine du gène de la transposase est donnée sous la séquence nucléique. Les motifs caractéristiques de la transposase mariner sont grisés pour les 2 HTH, souligné par un double trait pour le NLS, grisés pour les trois motifs de la triade catalytique avec chaque D indiqué en rouge et souligné. Les ITRs sont signalés par une flèche à chaque extrémité. La boîte TATA putative est entourée et les nucléotides indiqués en rouges. Les nombres notés à gauche correspondent respectivement aux nucléotides et aux acides aminés.

### ***Recherche des motifs autour de la triade catalytique (DD43D)***

Les trois principaux motifs conservés dans les transposases de végétaux : "HIDEKWF", "IQQDNARTH" et "NSPDL/MNLDLGFF" sont retrouvés mais avec quelques modifications. Dans la transposase *AaTnp*, ces motifs sont remplacés respectivement par "HVDEKWF", "IQQDGSTAH" et "QSPDTNIDLAF". Dans toutes les traductions des séquences complètes d'*Aacutmar1* et dans la séquence consensus *AaTnp*, la triade catalytique est toujours DD43D.

La structure spatiale putative de l'enzyme *AaTnp* a été recherchée à l'aide du programme de prédiction de structure Modeler. Ce programme compare une protéine à celles du même type dont la structure a été cristallographiée et qui sont décrites dans les banques de conformation de protéines. A ce jour, seule la transposase *marinerMos1* a été cristallographiée. Malgré la faible quantité de données présentes dans les banques pour les MLE, la comparaison des deux transposases a été tentée. Aucun résultat n'a été obtenu, ce qui peut s'expliquer par le fait qu'*AaTnp* diffère fortement de *Mos1*.

Les caractéristiques de la transposase *AaTnp* définies ci-dessus telles que la position des motifs HTH, NLS, la structure de la triade catalytique, etc. sont reprises dans le paragraphe suivant qui porte sur la comparaison d'*AaTnp* avec les transposases isolées de végétaux disponibles dans les banques de données.

#### **3.2.4. Comparaison de *ConsAacutmar1* avec des MLE de plantes disponibles dans les banques de données**

Une analyse par BlastX du transposon *ConsAacutmar1* a montré qu'il possède une homologie avec des MLE disponibles dans les banques de données. Les éléments complets qui lui sont les plus proches sont les MLE isolés d'Angiospermes parmi lesquels on peut citer les éléments : *Psmar1A* du petit pois *Pisum sativum* (34,2 % d'identité), *Costa* de la diatomée marine *Pseudo-nitzschia multiseries* (32,8 %), *Osmar1* du riz *Oryza sativa* (28,9 %), *Vulmar1* de la betterave *Beta vulgaris* (27,2 %), *Soymar1* du soja (31,8 %) et *Phmar1* du bambou *Phyllostachys edulis* (28,5 %) (Tab. 10).

Il est à noter que parmi ces éléments, *Costa* est un consensus, les autres étant des séquences mises en évidence dans les génomes par des manipulations de biologie moléculaire.

**Tableau 10 : Liste des transposases putatives dont la séquence protéique a été comparée avec *ConsAacutmar* et leur pourcentage d'homologie avec *ConsAacutmar*.**

Espèce	Elément	Homologie avec <i>ConsAacutmar</i>	Taille (aa)	Accession	Référence
<i>P. sativum</i> (pois)	<i>Psmar1A</i>	34,2	412	AY833549	Macas et al., 2005
<i>P. multiseriis</i> (diatomée)	<i>Costa</i>	32,8	503	JX013907	Non publiée
<i>G. max</i> (soja)	<i>Soymar1</i>	31,8	425	AF078934	Jarvick et Lark, 1998
<i>O. sativa</i> (riz)	<i>Osmar1</i>	28,9	576	AF172282	Tarchini et al., 2000
<i>P. edulis</i> (bambou)	<i>Phmar1</i>	28,5	499	HM581665	Zhou et al., 2011
<i>B. vulgaris</i> (betterave)	<i>Vulmar1</i>	27,2	519	AJ556159	Jacobs et al., 2004
<i>A. acutiuscula</i> (diatomée)	<i>ConsAacutmar1</i>	-	491		Présent travail

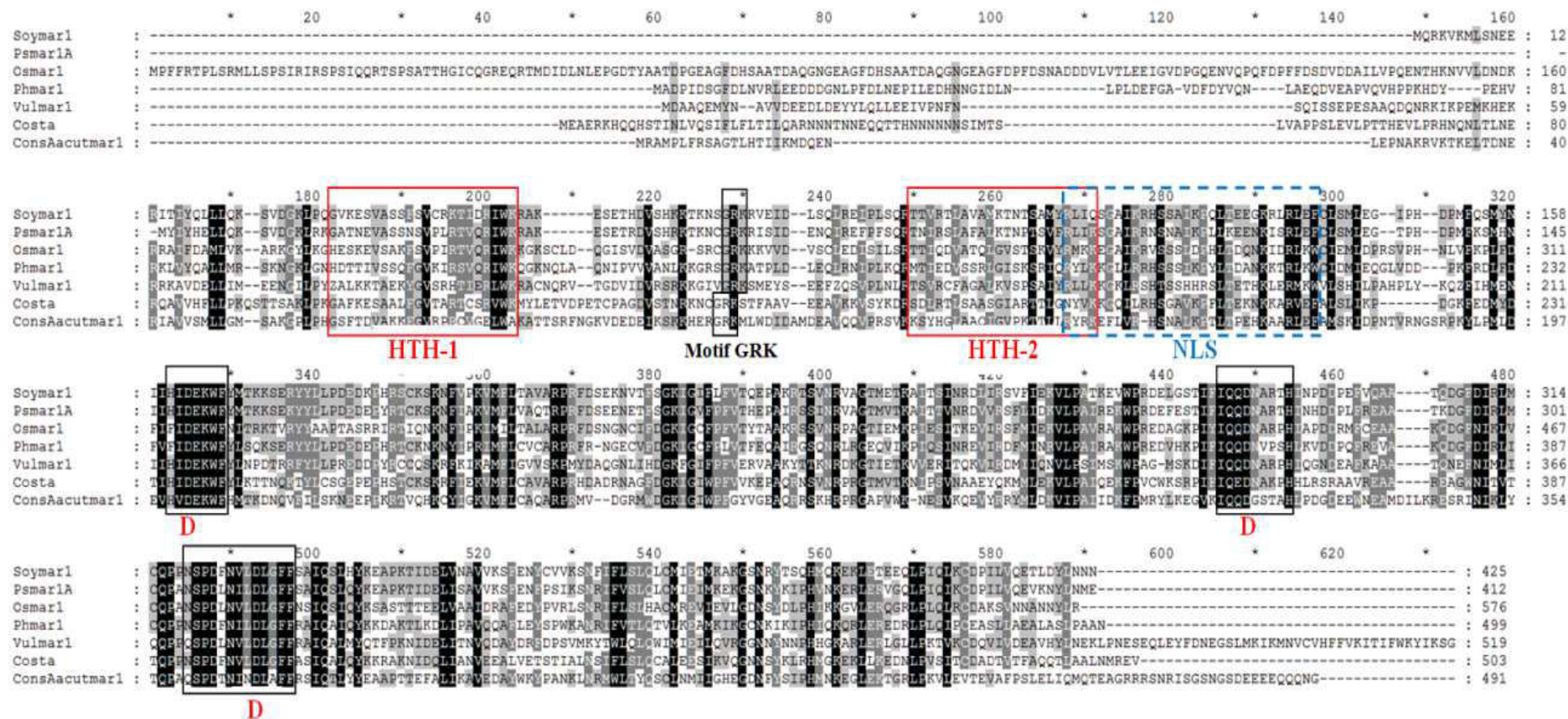
Au niveau de la triade catalytique, les trois blocs d'acidesaminés conservés autour de chaque acide aspartique sont également présents (Fig. 59). Les transposons isolés de végétaux sont caractérisés par une signature typique de la triade catalytique avec le motif DD39D, cette caractéristique n'est présente ni chez *ConsAacutmar1*, ni dans les séquences *Aacutmar1* prises individuellement qui possèdent toutes un motif DD43D comme mentionné précédemment. Toutes les transposases analysées ici présentent un ou deux motifs Hélice-Tour-Hélice (HTH), ceux-ci ont été détectés à l'aide du logiciel "HTH prediction" (Tab. 11). Chaque motif est constitué de 22 acides aminés. Les deux motifs sont espacés de 35 à 45 résidus avec, au centre de cette zone, un motif très conservé de trois acides aminés, GRK.

**Tableau 11 : Séquences des motifs Hélice-Tour-Hélice de transposases de plantes et de diatomées.**

Le score indique la probabilité que la protéine contienne un motif HTH. D1 : distance en acides aminés entre HTH-1 et le motif GRK. D2 : distance en acides aminés entre le motif GRK et HTH-2.

Élément	HTH1 (score - probabilité)	D1	Motif	D2	HTH 2 (score - probabilité)
<i>Osmar1</i>	HESKEVSAKFSVPRTVQRIWK (2.6 - Probabilité 25 %)	21	GRK	17	TTIQDVATQLGVSTSKVYRMKK (5.93 - Probabilité 100 %)
<i>Vulmar1</i>	YALKKTAKEYGVSRTTIERLWK (4.59 - Probabilité 100 %)	21	GRK	17	TSVRCFAGALKVSPSAIYRLK (4.60 - Possibilité 100 %)
<i>Soymar1</i>	GVKESVASSFSVCRKTIDRIWK (1.63 - Probabilité non significative)	18	GRK	17	TTVRTLAVAMKTNTSAMYRLIQ (2.95 - Probabilité 25 %)
<i>Costa</i>	GAFKESAALFGVTARTCSRVRWK (3.54 - Probabilité 70 %)	23	GRK	18	SDLRTLSAASGIARTTLGNYVK (1.31 - Probabilité non significative)
<i>Phmar1</i>	HDTTIVSSQFGVKIRSVQRIWK (1.19 - Probabilité significative)	22	GRK	17	MTIEDVSSRLGISKSRIQRYLK (5.96 - Probabilité 100 %)
<i>Psmar1A</i>	GATNEVASSNSVPLRTVQRIWK (0.2 - Probabilité non significative)	18	GRK	17	TNIRSLAFALKTNPTSVFRLIK (3.32 - Probabilité 50 %)
<i>ConsAacutmar1</i>	GSFTDVAKKFGVRPRCVGELWA (3.25 - Probabilité 50 %)	23	GRK	20	KSYHGLAAQIGVPKTTLLRYRK (2.02 - Probabilité non significative)





**Figure 58 : Alignement d'*AaTnp* avec les transposases d'autres éléments complets de la famille *mariner*.**

*ConsAacutmar1a* été traduit in silico en protéine nommée *AaTnp* puis aligné avec les transposases *Osmar1* (AF172282), *Vulmar1* (AJ556159), *Soymar1* (AF078934), *Psmar1A* (AY833549) *Costa* (JX013907), *Phmar1* (HM581665) en utilisant les logiciels ClustalW2 puis GeneDoc. Dans les séquences, les deux motifs HTH sont entourés en rouge et le NLS est entouré par des pointillés bleus. Les motifs conservés de la triade catalytique sont entourés en gris et les D sont positionnés en rouge.

Dans chaque transposase, un signal de localisation nucléaire (NLS) a été recherché à l'aide du logiciel NLS Mapper (Tab. 12 et Fig. 59). Comme mentionné précédemment, ce motif a été détecté dans *ConsAacutmar* avec un score moyen de 3,2. Parmi les transposases étudiées, la probabilité la plus élevée de présence d'un NLS a été détectée dans la transposase de l'élément *Soymar1* du soja (score de 8,5). Inversement, il a été noté que la prédiction du NLS dans la séquence de la transposase de l'élément *Costa* de la diatomée *P. multiseriis* donne un motif de NLS court de 12 acides aminés, "LTEKNKKARVEF". Ce motif présente des similarités avec d'autres NLS, ce qui explique sa détection par le logiciel avec une probabilité un peu inférieure à 50 % (score de 3). Les motifs NLS des transposases sont localisés à des positions similaires dans les différentes protéines. On peut remarquer également que dans chaque transposase, le NLS chevauche sur 4 acides aminés le deuxième HTH (Fig. 58).

**Tableau 12 : Identification des motifs NLS prédits dans les transposases de plantes et de diatomées.**

Élément	Motif NLS	Nombre aa	Score
<i>Osmar1</i>	RMKKEGAIKRVSSSLDPHLTDQNKIDRLKW	30	3,6
<i>Vulmar1</i>	RLLKKGKLRSHSTSSHRSLSLTHKLERMKW	30	4,1
<i>Soymar1</i>	RLIQSGAIKRHSSAIKPQLTEEGKRLRLEF	30	8,5
<i>Costa</i>	LTEKNKKARVEF	12	3,0
<i>Phmar1</i>	RYLKKGLLRHSSSIKPYLTDANKKTRLKW	30	4,4
<i>Psmar1A</i>	RLIKSGAIRRNSNAIKPLLKEENKISRLEF	30	2,5
<i>ConsAacutmar1</i>	RYRKEFLVRHSNALKPTLTPEHKAARLEF	29	3,2

Soymar1 : RLIQSGAIKRHSSAIKPQLTEEGKRLRLEF  
Psmar1A : RLIKSGAIRRNSNAIKPLLKEENKISRLEF  
Costa : NYVKKGQILRHSGAVKPFLLTEKNKKARVEF  
Aacutmar1 : RY-RKEFLVRHSNALKPTLTPEHKAARLEF  
Vulmar1 : RLLKKGKLRSHSTSSHRSLSLTHKLERMKW  
Phmar1 : RYLKKGLLRHSSSIKPYLTDANKKTRLKW  
Osmar1 : RMKKEGAIKRVSSSLDPHLTDQNKIDRLKW

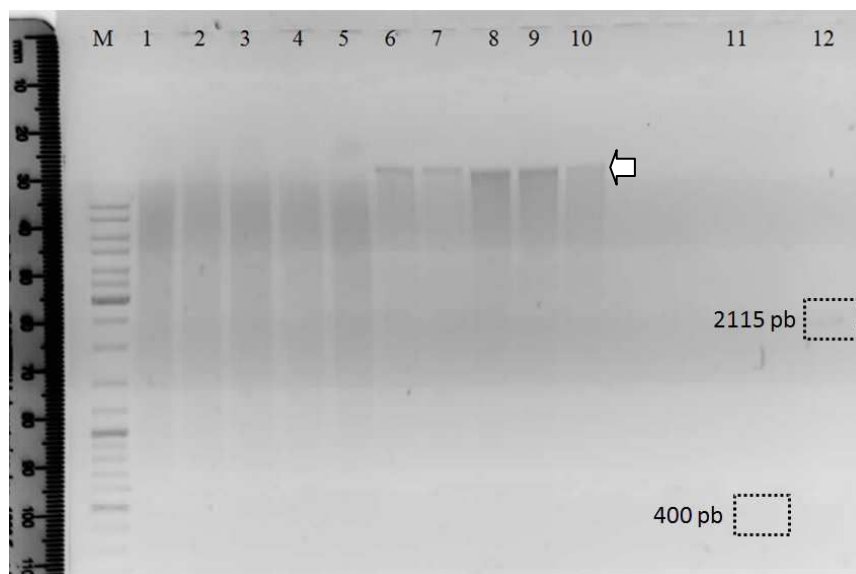
**Figure 59 : Alignement des motifs NLS des transposases de plantes et de diatomées.**

### **3.3. Estimation du nombre de copies d'*Aacutmar* dans le génome de *A. acutiuscula***

Après avoir caractérisé les éléments *Aacutmar*, nous nous sommes intéressés à leur nombre de copies dans le génome hôte. Cette estimation a été réalisée par deux méthodes : la première est qualitative et permet de visualiser l'envahissement du génome hôte par l'élément recherché, elle a été effectuée par Southern blot, la seconde est quantitative et permet d'estimer le nombre de copies, elle a été réalisée par PCR quantitative.

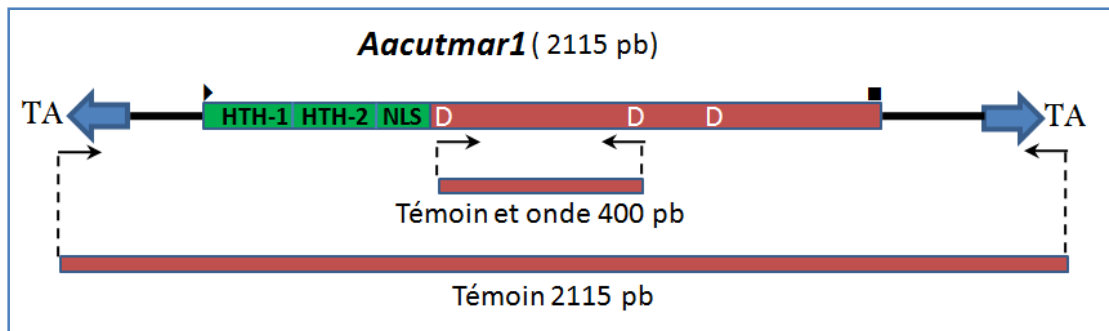
#### **3.3.1. Visualisation des copies d'*Aacutmar* par Southern blot**

Un Southern-blot a été réalisé à partir de cinq extractions d'ADNg digéré par *HindIII*. Outre les ADNg digérés, les ADNg non digérés ont également été transférés sur la membrane de nylon, (Fig. 60) ainsi que deux produits de PCR d'environ 400 pb et 2100 pb spécifiques de *Aacutmar1* et obtenus grâce aux amorces *MLE5A/MLE3A* et *AcutITR-5'/AcutITR-3'* (Fig. 61) respectivement, ces deux derniers servant de contrôles positifs de l'hybridation.



**Figure 60 : Migration sur gel d'agarose d'agarose 1 % des ADNg d'*A. acutiuscula*.**

M : marqueur de taille; puits de 1 à 5 : ADN génomiques d'*A. acutiuscula* digérés avec l'enzyme *HindIII*; puits de 6 à 10 : ADN génomiques d'*A. acutiuscula* non digérés; puits 11 : témoin d'hybridation n°1 produit de PCR de 400 pb obtenu avec les amorces *MLE5A* et *MLE3A* et puits 12 : témoin d'hybridation n°2 produit de PCR de 2100 pb obtenu avec les amorces *AcutITR-5'* et *AcutITR-3'*. Avant dépôt sur gel, les deux bandes constituant les témoins ont été purifiées par élution.

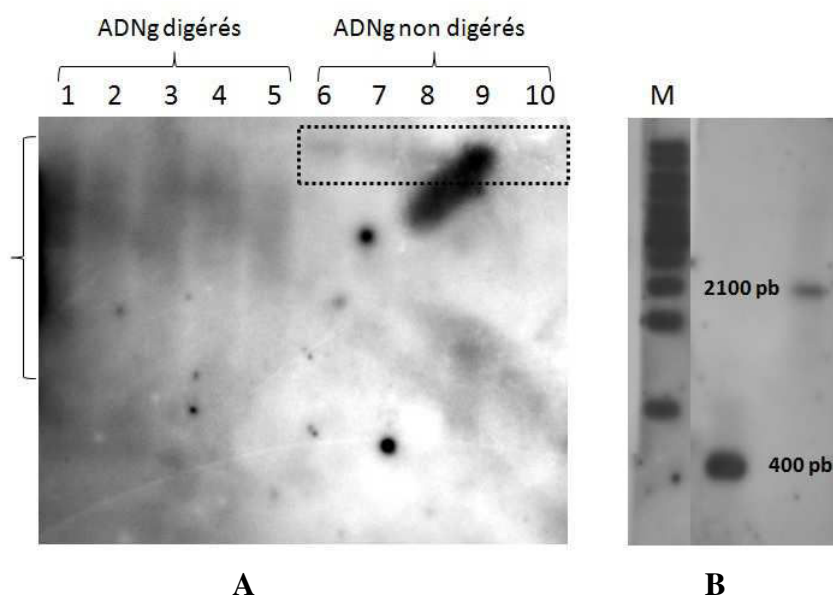


**Figure 61 : Schéma de la localisation de la sonde de 400 pb et des deux séquences témoins.**

Les flèches en noir représentent la localisation des amorces *MLE5A/MLE3A* et *AcutITR-5'/AcutITR-3'*

La figure 62-A montre les ADNg digérés et non digérés. Pour les ADNg non digérés, la présence de "smears" indique une légère dégradation du matériel génétique utilisé. Les deux produits de PCR-témoins déposés sur gel ne sont pas visibles car une très petite quantité (1 ng) a été déposée afin que le signal après hybridation ne soit pas trop important.

Afin de confirmer la présence des MLE chez *A. acutiuscula* et de visualiser l'envahissement du génome par *Aacutmar1*, la membrane de nylon issue du transfert des ADN a été hybridée avec une sonde d'environ 400 pb radiomarquée préparée à partir du produit de PCR obtenu avec les amorces *MLE5A* et *MLE3A* (Fig. 62-B).



**Figure 62 : Southern blot obtenu à partir d'un transfert d'ADNg et de témoins d'hybridation.**

**A)** ADNg digérés par HindIII et non digérés, **B)** témoins d'hybridation : produits de PCR d'environ 400 pb et 2115 pb amplifiés respectivement avec les amorces *MLE5A/MLE3A* et *AcutITR-5'/AcutITR-3'*.

La figure 62-B montre une hybridation des deux témoins confirmant bien l'accrochage des sondes sur leurs cibles. En ce qui concerne la présence d'*AacutmarI* dans le génome d'*A. acutiuscula*, celle-ci est confirmée par le signal, bien que faible, obtenu dans les puits 6 à 10. Parallèlement, sur les ADNg digérés, l'hybridation tout le long des "smears" (puits 1 à 5) indiquent un nombre de copies probablement élevé dans le génome-hôte. Dans tous les cas, les signaux obtenus pour les hybridations sont faibles mais visibles sur les autoradiogrammes.

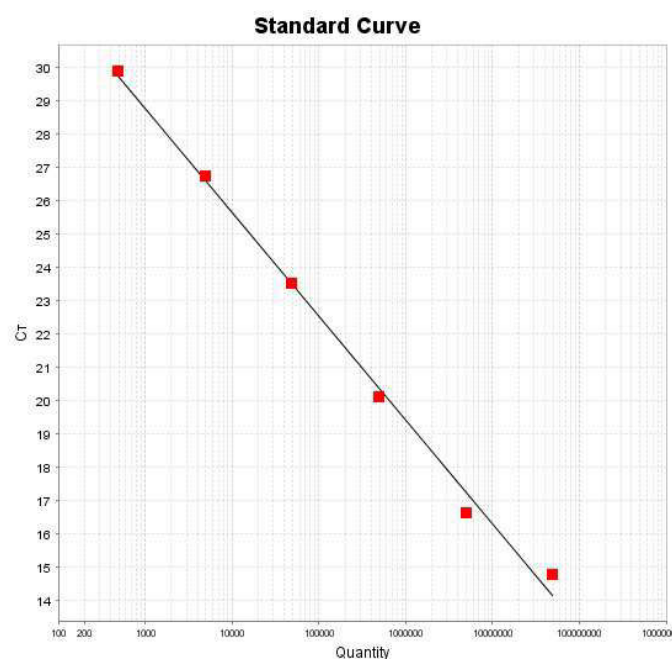
### **3.3.2 Estimation du nombre de copies par PCR quantitative**

Afin de quantifier le nombre de copies d'*AacutmarI* dans le génome d'*A. acutiuscula*, la technique de PCR en temps réel a été employée sur 4 des 5 échantillons d'ADNg utilisés en Southern-blot. L'une des extractions ayant été déposée en totalité sur gel pour le transfert, elle n'a pu servir pour la deuxième quantification. Parallèlement, la quantification a été réalisée sur une gamme étalon constituée par des dilutions d'un plasmide contenant un fragment du transposon *AacutmarI* de 96 pb préalablement amplifié avec les amorces *acutimarF1/acutimarR1*.

Pour la gamme étalon, des cultures bactériennes ont été réalisées afin d'extraire le plasmide contenant l'insert d'intérêt. La concentration en plasmides obtenue était d'environ 223 µg/mL soit 13,38 µg dans les 60 µL extraits. Sachant qu'un nucléotide a une masse moléculaire moyenne de 330 g/mol et que le plasmide contenant l'insert a une taille de 3096 pb et est double brin, sa masse moléculaire est de 2043360 g/mol. Le nombre de moles est alors de  $13,38 \cdot 10^{-6} / 2043360 = 6,54 \cdot 10^{-12}$  moles soit  $3,937 \cdot 10^{12}$  copies de plasmides dans les 60 µL extraits. 4,9 µL ont été utilisés, après une dilution de 5600 fois pour la PCR quantitative. Au final le nombre de copies de plasmides est estimé à  $5,75 \cdot 10^7$  dans chaque puits.

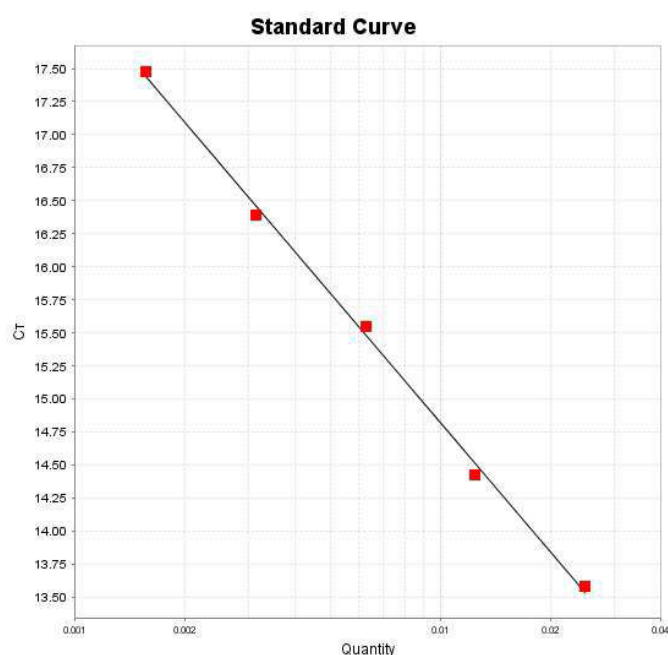
L'efficacité des amorces *acutimarF1/acutimarR1* sur la suspension de plasmides purifiés a été estimée à partir d'une gamme de dilutions allant de  $6 \cdot 10^7$  plasmides à 6 plasmides. L'efficacité obtenue était de 109,196 % et le  $R^2$  de 0,995 avec des amorces à une concentration finale de 0,8 µM (Fig. 63) avec une PAE de 1,065 ce qui montre une efficacité correcte et de bonnes conditions de Q-RT-PCR pour amplifier le fragment ciblé.





**Figure 63: Courbe standard obtenue avec les amorces *acutimarF1/R1* sur une gamme de concentrations en plasmides contenant un fragment du gène *Aacutmar1*.**

Parallèlement, l'efficacité de ces mêmes amorces a été déterminée sur 4,9  $\mu$ L d'ADNg à partir d'une gamme de dilutions (1/40, 1/80, 1/160, 1/320 et 1/640) réalisée sur les quatre ADNg différents. L'efficacité obtenue était de 103,639 % et le  $R^2$  de 0,998 avec des amorces à une concentration finale dans le milieu réactionnel de PCR en temps réel de 0,8  $\mu$ M (Fig. 64) avec une PAE de 1,04 montrant, comme pour le plasmide, l'efficacité des amorces à se fixer aux séquences recherchées.



**Figure 64 : Courbe standard obtenue sur ADNg avec les amorces ciblant le gène de la transposase d'*Aacutmar1*.**

Afin de pouvoir estimer le nombre de copies par génome, il s'est avéré nécessaire de dénombrer le nombre de cellules ayant été utilisé pour chaque extraction. Pour cela, le nombre de cellules en culture a été estimé à l'aide d'un hématimètre Neubauer. Après extraction des ADNg, chaque échantillon, d'un volume de 80  $\mu$ L, a été dosé par spectrophotométrie (Tab. 13).

La méthode utilisée pour calculer le nombre de copies de MLE *Aacutmar1* dans le génome est donc celle de la comparaison des Ct entre l'ADNg et le plasmide, dans chaque puits. Enfin, rapporté au volume extrait et au nombre de cellules dans la culture, le nombre de copies de MLE *Aacutmar1* dans les échantillons varie entre 11 à 29 (Tab. 13).

**Tableau 13 : Résultats de l'estimation du nombre de copies de MLE dans le génome d'*A. acutiuscula* NNC 216 par PCR en temps réel.**

	Nombre de cellules total dans la culture	Concentration en ng/μL	Ct	ΔCt	$2^{-\Delta Ct}$	Nombre de copies dans le puits	Nombre de copies dans l'extrait d'ADN	Nombre de copies par génome
ADN 1	186875000	694	20,7	-0,325	1,252	52078	2002519848	11
ADN 2	120625000	963	20,9	-0,134	1,098	45634	2258592143	19
ADN 3	131875000	2945	21,8	0,803	0,573	23825	3784896665	29
ADN 4	180750000	2955	21,7	0,657	0,634	26367	4427187039	24



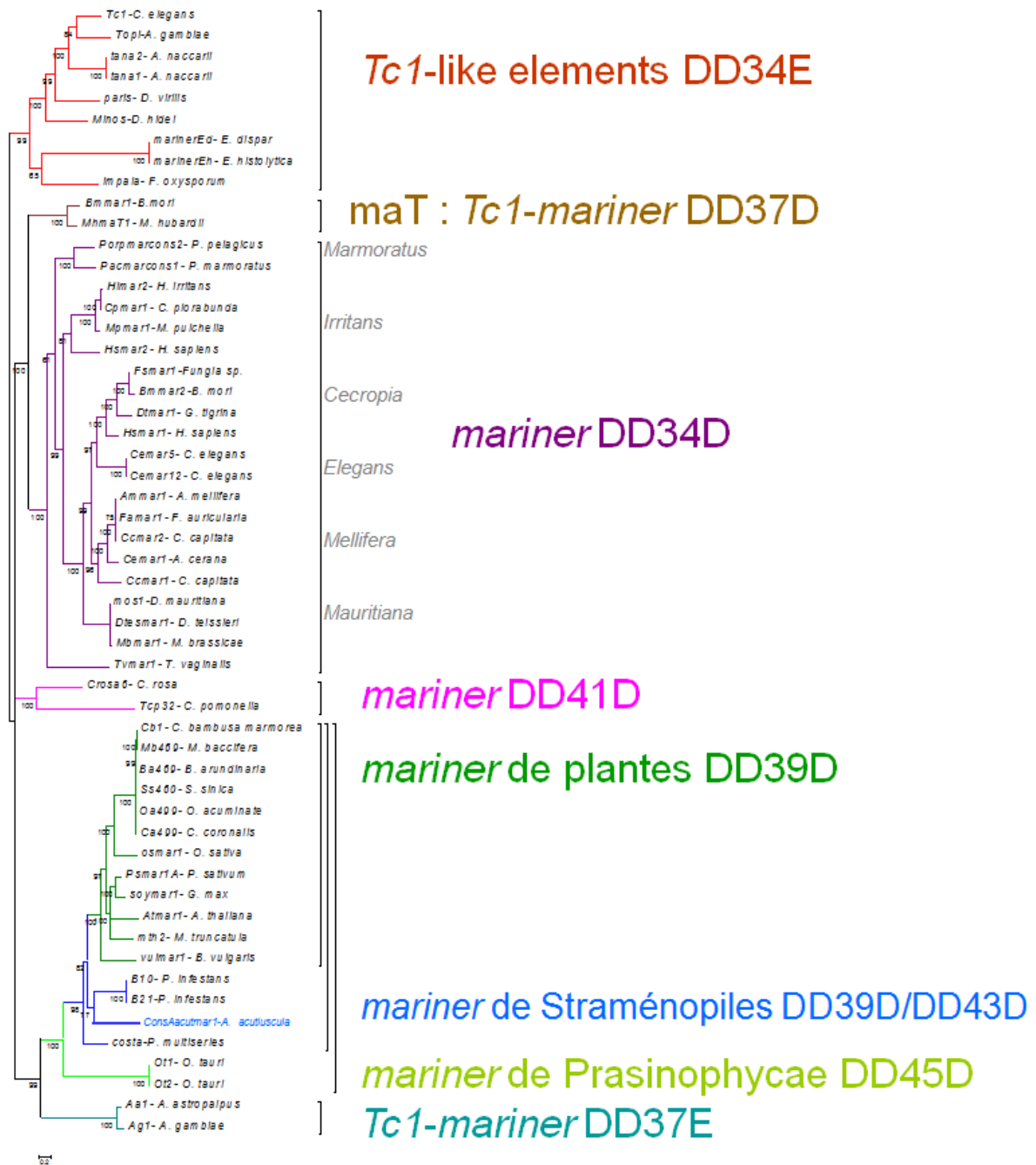
### **3.4. Relations phylogénétiques des MLE d'*A. acutiuscula* avec ceux des plantes et des autres Straménopiles présents dans les banques de données**

La position des MLE d'*A. acutiuscula* a été étudiée du point de vue phylogénétique par rapport à celle des éléments de la superfamille *Tc1-mariner* et plus particulièrement par rapport à celle des MLE de Straménopiles et de plantes Angiospermes. L'utilisation de l'Inférence Bayésienne a permis d'étudier les relations phylogénétiques présentes entre les MLE d'*A. acutiuscula* et ceux des sous-familles préalablement identifiées et ce, sur l'ensemble de la séquence du gène codant la transposase. L'arbre phylogénétique consensus présenté dans la figure 65 a été construit par le logiciel MrBayes et il a été enraciné avec des séquences de TLE (*Tc1*-like elements) présentant une triade catalytique de structure DD34E.

En excluant le groupe externe, l'arbre se divise en trois principaux groupes ou clades paraphylétiques soutenus par de fortes valeurs de probabilité postérieure, ce qui reflète une distance phylogénétique importante entre les clades et donc entre les éléments qui les composent.

- Le premier clade est constitué par les éléments de la sous famille *maT* qui possèdent une structure de triade catalytique de type DD37D et des éléments *mariner* et *mariner-like* qui ont une triade de type DD34D. Au sein de ce groupe, les sous-familles de MLE *marmoratus*, *irritans*, *cecropia*, *elegans*, *mellifera* et *mauritiana* ont été identifiées.
- Le second clade est constitué de deux séquences qui sont caractérisées par une structure de la triade putative de type DD41D, il correspond à la sous-famille *rosa*.
- Le troisième et dernier clade est soutenu par une valeur de probabilité postérieure de 99 et se divise en deux branches principales :

D'une part, les séquences dont la structure protéique putative correspond à une triade de type DD37E et d'autre part, celles ayant une structure de la triade catalytique putative DD39D, DD43D et DD45D. Cette branche est supportée par une valeur de probabilité postérieure de 100 et regroupe des séquences isolées d'espèces appartenant aux lignées des plantes et des Straménopiles, c'est-à-dire des organismes hôtes possédant des chloroplastes (plantes, algues vertes, diatomées) ou en ayant possédés (oomycètes). La majorité de ces séquences provient de génomes de plantes. C'est dans ce dernier clade qu'est positionnée la séquence consensus MLE d'*A. acutiuscula* présentant la structure DD43D.



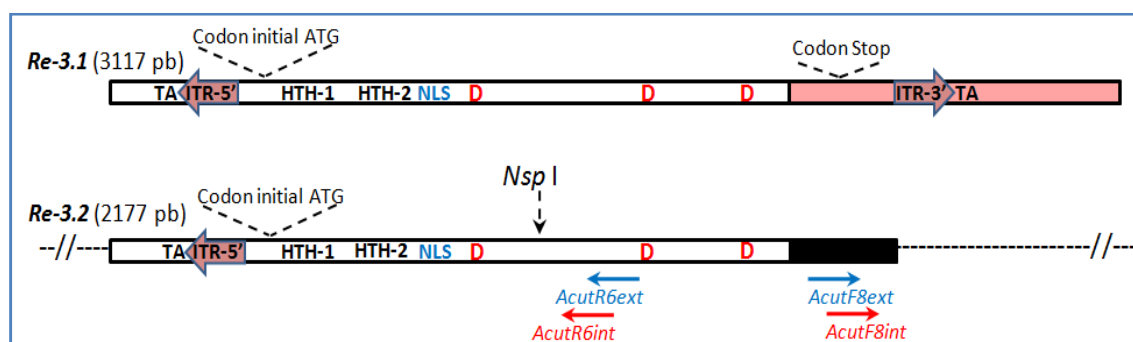
**Figure 65 : Relations phylogénétiques entre les transposases complètes des MLE de Straménopiles et les MLE des plantes angiospermes terrestres.**

L'arbre a été généré par comparaison des séquences de MLE de la diatomée *A. acutiuscula* obtenue (en bleue) dans le présent travail avec les MLE de Straménopiles et de plantes terrestres extraits des banques de données. Les relations phylogénétiques ont été établies par référence avec les autres séquences de transposases complètes de la superfamille *Tc1-mariner*. En rouge est indiqué le groupe des *Tc1*-like elements représentant le groupe externe de cette analyse. Chaque groupe de la superfamille *Tc1-mariner* est représenté par une couleur différente et les sousfamilles précédemment identifiées sont indiquées en gris. Les MLE de plantes sont représentés en vert tandis que les MLE de Straménopiles sont indiqués en bleu. Les noms de chacun des éléments sont suivis du nom de l'espèce au sein de laquelle l'élément a été mis en évidence. L'analyse phylogénétique a été réalisée grâce au logiciel Mr Bayes avec la méthode d'inférence Bayésienne et les probabilités postérieures sont indiquées lorsqu'elles sont supérieures à 50.

Il est à noter que les éléments DD39D identifiés chez les plantes sont proches des éléments DD39D identifiés chez une diatomée (*Costa* identifié *in silico* chez *P. multiseries*) et des oomycètes (*B10* et *B21* identifiés *in silico* chez *P. infestans*). Nous pouvons également noter que la séquence MLE consensus d'*A. acutiuscula* se positionne au milieu de ces éléments bien qu'elle présente une structure de la triade catalytique putative de type DD43D. Les MLE isolés chez la microalgue verte *O. tauri* (Prasinophyceae) se positionnent sur une branche supportée par une valeur de probabilité postérieure de 100 et qui diverge très tôt par rapport au reste du groupe indiquant que ces séquences ont probablement évolué de manière plus rapide que celles du reste du groupe. Il apparaît également que les éléments de *O. tauri* ont probablement accumulé plus de mutations comme le suggère la longueur de la branche qui les porte. Cette analyse permet de montrer que les MLE d'*A. acutiusculas* ont dans le groupe des MLE des Straménopiles et des plantes, indiquant clairement que ces éléments sont phylogénétiquement apparentés.

#### 4. Allongement de l'extrémité 3' de l'élément *Re-3.2*

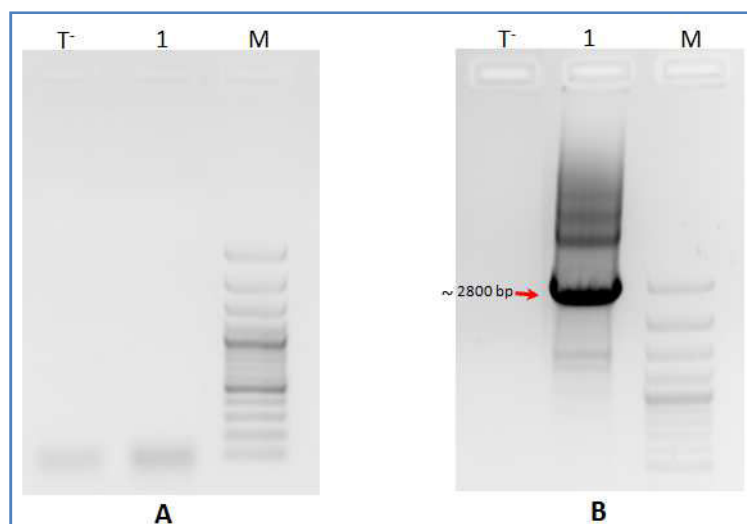
Dans la séquence reconstruite *Re-3.1*, nous avons pu identifier un transposon complet avec ses deux ITR mais dans le fragment *Re-3.2* qui est un peu plus court que l'élément *Re-3.1*, seul l'ITR-5' a été identifié. Nous avons donc poursuivi l'allongement de l'extrémité 3' de la séquence *Re-3.2* pour rechercher l'ITR-3'. Les deux séquences reconstruites *Re-3.1* et *-3.2* partagent une séquence identique de 1782 pb en région 5' mais sont différentes en région 3'. Pour cela, l'enzyme *NspI* qui coupe à l'intérieur de *Re-3.2* a été choisie et deux couples d'amorces chevauchantes ont été définis, les amorces sens *AcutR6* (*ext* et *int*) situées dans la partie commune des *Re-3* et les amorces antisens *AcutF8* (*ext* et *int*) situées dans la zone spécifique de la séquence *Re-3.2* (Fig. 66).



**Figure 66 : Schéma permettant de localiser les amorces ayant servi à l'allongement de l'extrémité 3' de l'élément *Re-3.2*.**

L'enzyme *NspI* qui coupe à l'intérieur de la séquence *Re-3.2* au niveau du domaine catalytique a été utilisée. Les flèches en bleu et en rouge représentent les deux couples d'amorces chevauchantes utilisés. Les amorces *AcutR6ext* et *AcutF8ext* (pour externe) ont été utilisées pour l'iPCR n°1 et les amorces *AcutR6int* et *AcutF8int* (pour interne) pour l'iPCR n°2. La ligne en pointillés représente les régions bordant l'élément *Re-3.2*.

Après l'iPCR, un fragment d'environ 2800 pb a été obtenu (Fig. 67). Il a été cloné et deux clones ont été séquencés. Ces derniers ont été nommés *NspIR6F8-1* et *NspIR6F8-2*, ils sont similaires à 99,3 %, de ce fait, seul, le clone *NspIR6F8-1* a été utilisé pour l'analyse. Il a été séparé en deux parties, *NspIR6F8-5'* et *NspIR6F8-3'*, au niveau du site de restriction de l'enzyme *NspI*, puis le fragment *NspIR6F8-3'* (1912 pb) a été aligné avec la région 3' de la séquence reconstruite *Re-3.2* (Fig. 68-A).



**Figure 67 : Migration sur gel d'agarose des produits d'amplification du fragment bordant l'extrémité 3' de la séquence *Re-3.2*.**

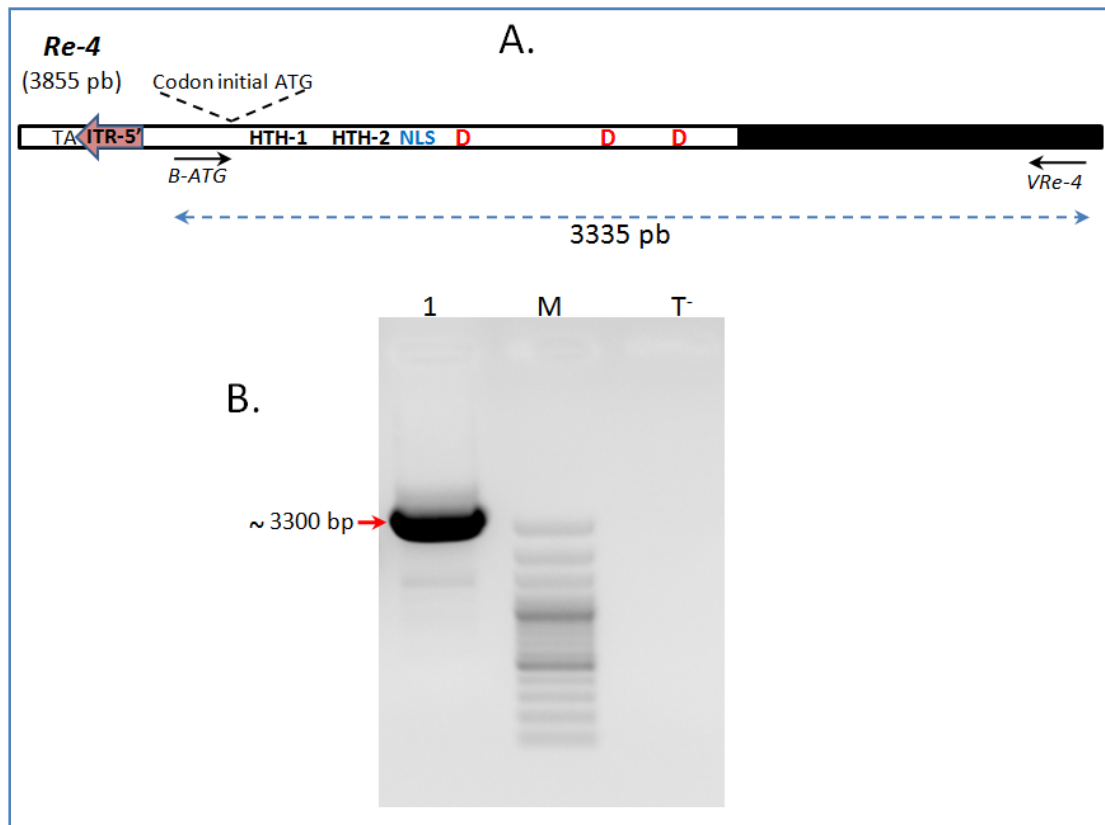
**A** - Puits 1 : produit de la première iPCR utilisant comme matrice l'ADNg digéré par l'enzyme *NspI* et le couple d'amorces externes *AcutR6ext* et *AcutF8ext*. **B** - Puits 1 : produit de la seconde iPCR utilisant comme matrice le produit de l'iPCR précédente et le couple d'amorces internes *AcutR6int* et *AcutF8int*. M : marqueur de taille GeneRuler 100 pb. T<sup>-</sup> : témoin négatif (iPCR réalisée sans ADNg).

L'assemblage des séquences *Re-3.2* et *NspIR6F8-3'* obtenues par iPCR a permis de reconstruire un nouvel élément nommé *Re-4* pour reconstruit n°4. Celui-ci mesure 3855 pb (Fig. 68-B). Des amorces ont été définies sur ce dernier pour vérifier sa présence effective dans le génome d'*A. acutiuscula* (Fig. 69-A).

L'amplification effectuée sur l'ADNg avec les amorces *B-ATG* et *VRe-4* a donné un fragment d'environ 3300 pb (Fig. 69-B). Après séquençage, un clone de 3303 pb nommé *BATG/VRe-4* a été obtenu et analysé. Sa comparaison par alignement avec la séquence reconstruite *Re-4* indique une identité de 94,88 % au niveau nucléique (alignement présenté dans l'annexe 5 page 234). L'ITR3' est absent de cette séquence qui est tronquée en région 3', ce nouvel élément est nommé *Aacutmar4*.

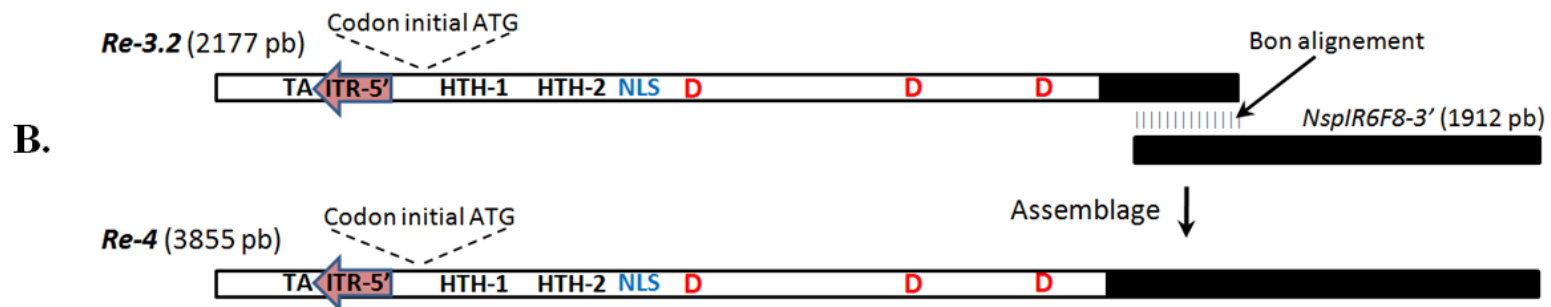
La comparaison par alignement de *ConsAacutmar1* et d'*Aacutmar4* montre une similarité de 86 % sur les 1474 premières paires de bases (Fig. 70). *Aacutmar4* a perdu la partie finale de 144 pb située de la fin du gène de transposase à l'extrémité de l'ITR-3'. *Aacutmar4* possède des mutations ponctuelles générant plusieurs codons stop et un décalage du cadre de lecture. *Aacutmar4* présente trois délétions importantes: une grande délétion de 24 pb "TCAAAACAAAAGAACTCACTGATA" (positions 271 à 294), une

de 3 pb "GTC" en 745-747 et une de 8 pb "AGATGGCC" (positions 926 à 933) (Fig. 70). Ces délétions sont entourées de courtes répétitions directes (micro-homologies de 4 à 6 pb) autour du point de délétion. Les micro-homologies: "AAGA" pour la délétion de 24 pb ; "GAATGT" pour la délétion de 3 pb et "GGATGG" pour la délétion de 8 pb.



**Figure 68 : A) Schéma de l'élément reconstruit *Re-4* montrant la localisation des amorces définies pour vérifier sa présence dans le génome d'*A. acutiuscula*. B) Amplification du fragment de vérification de la présence de *Re-4*.**

En résumé, les éléments MLE *Aacutmar* peuvent être séparés en quatre types distincts, les éléments complets *Aacutmar1* sous une forme presque exclusivement intacte; les éléments délétés *Aacutmar2* qui ont perdu une grande partie interne du gène de la transposase ; les éléments délétés *Aacutmar3* ont perdu une grande partie qui a été remplacée par une séquence ne correspondant pas à un ET et les éléments tronqués en région 3' *Aacutmar4* qui sont probablement immobiles du fait de la perte de leur ITR-3'. Les trois premiers types contenant les ITR peuvent quant à eux être mobilisés par une transposase *AaTnp*.



**Figure 69 : A) Schéma de l'alignement du fragment *NspIR6F8-3'* avec la séquence *Re-3.2*. B) L'assemblage du fragment *NspIR6F8-3'* avec *Re-3.2* a permis de former la séquence *Re-4*.**

Dans celle-ci ont été détectés depuis la région 5' vers 3' : le site cible TA, l'ITR5', les motifs HTH 1 et 2 (écrits en noir), le NLS (en bleu), les trois D de la triade catalytique (en rouge). L'ITR3' n'a, quant à lui, pas pu être identifié, il est donc absent de cette séquence.

```

      *           20           *           40           *           60           *
Aacutmar4 : CTCCCTCTGGTCCAAAATTAAACATCACAATCGGCTCTCTCTAAATAAAAAATTTGAATATCACCCCTTCCTAT : 74
ConsAacutmar1 : CCCCTCTGGTCCAAAATTAAACATCACAAATCGGCTCTCTCTAAATAAAAAATTTGAATATCACCCCTTCCTAT : 74
                  C CCCTCTGGTCCAAAATTAAACATCAC ATCGGCTCT TCTTAAATAAAAAATTT GAATATCAC CCT CCTAT

      80           *           100           *           120           *           140
Aacutmar4 : TAGTACGTTATATCATTGCAATCTACTCAAAAAGAGCTACATATCCATGTATATATTGAATATTCAATTTTGAA : 148
ConsAacutmar1 : TAGTACGTTATATCATTGCAATCTACTCAAAAAGAGCTACATATCCATGTATATATTGAATATTCAATTTTGAA : 148
                  TAGTACGTTATATCATTG AATCTACTCAAAAAGAGCTACATATCCATGTATATATTGAATATT AATTTTGAA

      *           160           *           180           *           200           *           220
Aacutmar4 : CAGAAGGCAATGAAAAGGCGGGCTAAAAGGCATGCGTGCAAGGTTCTTTTGGATCACCGGGAACCTTGCACA : 222
ConsAacutmar1 : CAGAAGGCAATGAAAAGGCGGGCTAAAAGGCATGCGTGCAAGGTTCTTTTGGATCACCGGGAACCTTGCACA : 222
                  CAGAAGGCAATGAAAAGGC GGCTAAAAGGCATGCGTGCA A G CTT TT GATCA C GGAAC TGCACA

      *           240           *           260           *           280           *
Aacutmar4 : CAATTAACAAAATGGACCAGGAGAACCTTGAACCAAAACCAAGAGGG-----AT : 272
ConsAacutmar1 : CAATTAACAAAATGGACCAGGAGAACCTTGAACCAAAACCAAGAGGGTCAAAAACAAAAGAACTCAGTATAT : 296
                  CAAT A CAAAATGGACCAGGAGAACCTTGAACCAAA CCAA AGGG Délétion de 24 pb AT

      300           *           320           *           340           *           360           *
Aacutmar4 : GAAAGATTGCAAGTGGTTCCATGCTGTTAGGAATGTCTGCAAAAGGCCCTCTTCTCATGCTAGTTTACAGA : 346
ConsAacutmar1 : GAAAGATTGCAAGTGGTTCCATGCTGTTAGGAATGTCTGCAAAAGGCCCTCTTCTCATGCTAGTTTACAGA : 370
                  GAAAG ATTGCAAGTGGT TCCATG TGTAGGAATGTG CAAAAGGCCCTCTT CTATGCTAGTTT ACAGA

      380           *           400           *           420           *           440
Aacutmar4 : CGTGGCTAAAAAATTGGAGTCTGACCGTGCTGTGTGAGAGCTGTGGGCCAAGGCTACAACCTCTGTTTCA : 420
ConsAacutmar1 : TGTGGCTAAAAAATTGGAGTCTGACCGTGCTGTGTGAGAGCTGTGGGCCAAGGCTACAACCTCTGTTTCA : 444
                  GTGGCTAAAAAATTGGAGTG GACCG GCTGTGT GAGAGCTGTGGGCCAAGGCTACAACCTC T T CA

      *           460           *           480           *           500           *           520
Aacutmar4 : ATGGCAAAGTGATGAAGATGAGCTCAAATCAAAAGAACATGAGAGAGGAAGAAAGATGCTCTGCGATATTCAT : 494
ConsAacutmar1 : ATGGCAAAGTGATGAAGATGAGCTCAAATCAAAAGAACATGAGAGAGGAAGAAAGATGCTGTTGGATATTCAT : 518
                  ATGGCAAAGT GATGAAGATGAG TCAAATCAAAAGAA CATGAGAGAGGAAGAAAGATG TGT GGATAT GAT

      20           *           540           *           560           *           580           *
Aacutmar4 : GCAATGATGAAGCTGTGCAACAAGTTCCTTGCTCAGTGAAGAAGTCTTACCATGGTTGGCAGCACAGATAGG : 568
ConsAacutmar1 : GCAATGATGAAGCTGTGCAACAAGTTCCTTGCTCAGTGAAGAAGTCTTACCATGGTTGGCAGCACAGATAGG : 592
                  GCAAT GATGAAGCTGTGCAACAAGTTCCT GCTCAGTGAAGAAGTCTTACCATGG TTGGCAGCACAGATAGG

      600           *           620           *           640           *           660
Aacutmar4 : AGTCCCCAAGACAACCCCTGTTGCAATACAGAAAGGATTTCTTGTAGACATATCAATGCCCTCAAGCCGACAC : 642
ConsAacutmar1 : AGTCCCCAAGACAACCCCTGTTGCAATACAGAAAGGATTTCTTGTAGACACAGCAATGCCCTCAAGCCGACAC : 666
                  AGTCCCCAAGACAACCCCTGTTGC ATACAGAAA GA TTC TTGTAGACA A CAATGCCCT AAGCCGACAC

      *           680           *           700           *           720           *           740
Aacutmar4 : TCACTCCTGAACACATAGTAGCAAGACTGAGTTTTCATGAGTAAAGTAGACCCCAATATGGTAAGGAATGGA : 716
ConsAacutmar1 : TCACTCCTGAACACATAGTAGCAAGACTGAGTTTTCATGAGTAAAGTAGACCCCAATATGGTAAGGAATGGA : 740
                  TCACTCCTGAACACA AG AGCAAGACT GAGTTT C ATGAG AAA TAGACCC AATA GGTAAGGAATGGA

      *           760           *           780           *           800           *
Aacutmar4 : AGTCC---CCAAAGTACTTACCAATGTTAGATGAGGTTTCATGTTGATGAGAAGTGGTTTACATGACAAAGACAA : 787
ConsAacutmar1 : AGTCCGTTCCAAAGTACTTACCAATGTTAGATGAGGTTTCATGTTGATGAGAAGTGGTTTACATGACAAAGACAA : 814
                  AGTC C AAGTACTTACCAATGTTAGA GAGGTTTCATGTTGATGAGAAGTGGTT CACATGACAAA GACAA
Délétion de 3 pb

      820           *           840           *           860           *           880
Aacutmar4 : TCAAGTTTTCATCTGTGCAAGAGCCGAGAACCCCAAGAAACTGTTTCAACACAAAGGCTATACATTGGCAAAG : 861
ConsAacutmar1 : TCAAGTTTTCATCTGTGCAAGAGCCGAGAACCCCAAGAAACTGTTTCAACACAAAGGCTATACATTGGCAAAG : 888
                  TCAAGT TTCATC TGTC AAGA CGAG AAC CCCAA A AACTGT CA CACAA G TACATTGGCAAAG

      *           900           *           920           *           940           *           960
Aacutmar4 : TGAATCTTGTGTGTCACAAGCAAGCCCAAGCATCGT-----GAATCTGGGATGGCAAGCTTGGCATTTGG : 926
ConsAacutmar1 : TGAATCTTGTGTGTCACAAGCAAGCCCAAGCATCGTGAATCTGGGATGGCAAGCTTGGCATTTGG : 962
                  TGAATCTTGTGTGTCACAAGC G CCAAG ATGGT GAATCTGGGATGG AAG TTGGCAT TGG
Délétion de 8p

      *           980           *           1000           *           1020
Aacutmar4 : CCCTTTGGATATGTGGGAGAGGCAAGCAAAATATCAAAAGTATGAGACCAAGCTGGAGCCCTGTTTGGCAAAAGGA : 1000
ConsAacutmar1 : CCCTTTGGATATGTGGGAGAGGCAAGCAAGCAAAATATCAAAAGTATGAGACCAAGCTGGAGCCCTGTTTGGCAAAAGGA : 1036
                  CC TTTGGATATGT GGA A GC CA A A ATCAAA A AGACC GGAGC CC GTTTGG A AA GA

      1040           *           1060           *           1080           *           1100
Aacutmar4 : GAGTGTGATAGAGATCTTACCAATATATGCTTGACAAAGTATTTTCCGTCATCTTGGCAAGTTTCTTA : 1074
ConsAacutmar1 : GAGGTTTAAACCAAGAGGTTTACCGCGATACATCTTGACAAAGTATTTTCCGTCATCTTGGCAAGTTTCTTA : 1110
                  GAG GTT A AGA TAC C ATA ATG T GACAA GT ATT C GCCAT T GA AAGTT C A

      1120           *           1140           *           1160           *           1180
Aacutmar4 : TGGGTATCTCTGAAATGCGGGTGTGAAATCCAGCAAGAGGGTCAAAATCCCATATTCTGTGCAATGATGAA : 1148
ConsAacutmar1 : TGGGATATCTCTGAAATGCGGGTGTGAAATCCCAAGAGTGGATCAACCGCCCATTTGCTGTGATGATGAA : 1181
                  TG G AT CT TGAAA G GGG GT AA ATCCA CAAGA GG CAA CCCAT T CTGA GATGAA

      *           1200           *           1220           *           1240           *           12
Aacutmar4 : GAATGGTTAGATGCGATGCAAGCAATTTACG-ATGAGTACAGCTGAATATTCAGATATATACCAACCGGCTC : 1221
ConsAacutmar1 : GAATGGTTAGATGCGATGCAAGCAATTTACG-ATGAGTACAGCTGAATATTCAGATATATACCAACCGGCTC : 1254
                  GA TGG GA GC ATG A A T A TGAG CA T AA AT AG T TA AC CAACCGGCTC

```



```

      60          *          1280          *          1300          *          1320          *
Aacutmar4 : AATCACCAGACACAAACATCAATGATTGGCATTCTTTAGAGCATCCAAATGCTTTATTATGAAGCGCACC : 1295
ConsAacutmar1 : AATCACCAGACACAAACATCAATGATTGGCATTCTTTAGAGCATCCAAATGCTTTATTATGAAGCGCACC : 1328
      AATCACC GACACAAACATCAATGATTGGCATTCTT AG AGCATCCAA CTTTATTATGAAGC GCACC

      1340          *          1360          *          1380          *          1400
Aacutmar4 : ACAACCGACTTTGCTCTCATCAAGAGAGTGGAGATGCATATTGGAAGTATCCAGTTACAAGCTCAATCAGAT : 1369
ConsAacutmar1 : ACAACCGACTTTGCTCTCATCAAGAGAGTGGAGATGCATATTGGAAGTATCCAGTTACAAGCTCAATCAGAT : 1402
      ACAAC CGA TTGC CTCATCAAG AGT GAGGATGCATA TGGAGTATCCAG AACAAGCTCAATC AT

      *          1420          *          1440          *          1460          *          1480
Aacutmar4 : GTG-TTCACTTACCAAAGCTGCTTAACATGATCATTGACCGAGGGAGACAACCTTTTACTCAATTCCTCCCA : 1442
ConsAacutmar1 : GTGCTTCACTTACCAAAGCTGCTTAACATGATCATTGACCGAGGGAGACAACCTTTTACTCAATTCCTCCCA : 1476
      GTG T ACTTACCAAAGCTG CT AACATGATCATTG AC CGAGGGAGACAACCTTTTAC ATTCCCTC CA

      *          1500          *          1520          *          1540          *
Aacutmar4 : TGAACAAAGAGGGCTTGGAAAAACCGGCCGATGGCCAAAAGTCTTGGGAAGTACCGAAGTTGCCCTTTCCAAGT : 1474
ConsAacutmar1 : TGAACAAAGAGGGCTTGGAAAAACCGGCCGATGGCCAAAAGTCTTGGGAAGTACCGAAGTTGCCCTTTCCAAGT : 1550
      TGAACAAAGAGGGCTTGGAAAA ACCGGCCGA
      Position tronquée d'Aacutmar4

      1560          *          1580          *          1600          *          1620
Aacutmar4 : ----- : -
ConsAacutmar1 : CTTGAACTGATCCAAATGCAGACAGAGGCAGGAAGAAGAAGTAATCGTATTTCTGGCTCTAATGGTTCTGA : 1624

      *          1640          *          1660          *          1680          *          1700
Aacutmar4 : ----- : -
ConsAacutmar1 : TGAAGAAGAAGAACAAACAGAAATGGTTAGCAGTTGATCGAACAGGATGACCGAACAAGATGACCAAAACCAAG : 1698

      *          1720          *          1740          *          1760          *
Aacutmar4 : ----- : -
ConsAacutmar1 : ATGACCGAATCAAGGCCAAAACAAGGCCAATCCCCACAAAAGCACACCGATAGCACCACAATAACACAGAGAA : 1772

      1780          *          1800          *          1820          *          1840          *
Aacutmar4 : ----- : -
ConsAacutmar1 : CAAGAAGTTACACCGTATTGAGTTAAACACCACAAATAACAAAGAAATCCCATGGTCTTTGATCTAAAACAGA : 1846

      1860          *          1880          *          1900          *          1920
Aacutmar4 : ----- : -
ConsAacutmar1 : GAGCTGCTGCTGGAATCAAGTTTGTGATCGAAAATGAGCTCTGCAGCATGCTGCTGAGGAGTTCAGCGTGGAAGA : 1920

      *          1940          *          1960          *          1980          *          20
Aacutmar4 : ----- : -
ConsAacutmar1 : ACCTTCGCTCCGTGGCGCCAAAGAGAACCCAAAAGGAGAGATCTGATTGGCTAAAAATGTCTCCACGATGTTCT : 1994

      00          *          2020          *          2040          *          2060          *
Aacutmar4 : ----- : -
ConsAacutmar1 : TTCGCGCCACGTACTTATGAATGTTCCACGCTAACACAGCCCTGTAGAATTTGTACCAGTTCTACGTGGGTCTT : 2068

      2080          *          2100          *
Aacutmar4 : ----- : -
ConsAacutmar1 : GTTGTGCTCGTACACGGGACTGATGAATAATCTGGACCAGAGGGAG : 2114

```

**Figure 70 : Alignement de *ConsAacutmar1* et d'*Aacutmar4*.**

## **5. Discussion**

### **5.1. Présence des MLE dans le génome d'*A. acutiuscula* et obtention d'un élément complet**

Dans le présent travail, nous avons recherché les éléments transposables de type *mariner* dans le génome de la microalgue marine *A. acutiuscula*. Ces transposons ont été choisis car ils sont très répandus chez les eucaryotes. S'ils ont été bien étudiés chez les animaux, notamment les insectes, chez les plantes, ou encore chez l'homme, ils restent, à ce jour, peu décrits chez les microalgues (Robertson et al., 2002 ; Feschotte et Wessler, 2002 ; Plasterk et van Luenen, 2002 ; Robertson et Zumpano, 1997 ; Robertson et Martos, 1997). Pour ces dernières, on peut citer le dépôt récent dans les banques de données de l'élément complet *Costa* découvert *in silico* dans le génome séquencé de la diatomée marine *P. multiseriata* (n° d'accès JX013907, non publié) et des travaux ayant conduit à l'obtention de séquences partielles de MLE chez 10 espèces de diatomées dont *A. acutiuscula* (Hermann et al., soumis) et ce, grâce aux amorces dégénérées définies par Feschotte et Wessler en 2002. Ces amorces bordent un fragment interne du gène de la transposase d'environ 400 pb, situé entre les deux premiers résidus acides aspartiques de la triade catalytique et couvrant environ 130 acides aminés du cadre de lecture de la protéine. Chez de nombreuses plantes, l'utilisation de ces amorces a généré une ou plusieurs séquences partielles correspondant au gène de la transposase de MLE (Feschotte et Wessler, 2002, Jacobs et al., 2004 ; Zhou et al., 2010). Chez *A. acutiuscula*, ces mêmes amorces ont permis de détecter des fragments partiels de MLE dans le génome hôte (Thèse D. Hermann, 2011 et présent travail). Une fois les MLE détectés, nous nous sommes attachés à obtenir l'élément complet. Pour cela, une stratégie de PCR inverse a été mise en place afin d'amplifier les régions bordant la séquence initiale de 400 pb et permettre, en marchant de proche en proche le long du fragment, d'obtenir un élément complet avec ses ITR. Après une série d'amplifications par iPCR, nous avons pu obtenir par assemblage *in silico* un élément pleine taille nommé *Aacutmar1-re* et mesurant environ 2130 pb. *Aacutmar1-re* étant une reconstruction informatique de l'élément, sa présence dans le génome d'*A. acutiuscula* devait encore être confirmée. Pour cela les ITR ont servi de matrice pour définir des amorces susceptibles de permettre l'amplification par PCR d'un transposon *Aacutmar* complet à partir de l'ADNg d'*A. acutiuscula*.

## **5.2. Amplification des MLE dans le génome d'*A. acutiuscula* avec les amorces d'ITR**

A l'aide des amorces définies sur les ITR, trois types de MLE ont été isolés : *Aacutmar1*, *Aacutmar2* et *Aacutmar3*. *Aacutmar1* est l'élément pleine taille alors que *Aacutmar2* et *3* sont des éléments tronqués.

Pour *Aacutmar1*, au total, 12 séquences différentes ont été amplifiées, elles ont une taille comprise entre 2094 et 2132 pb et des pourcentages d'homologie qui varient entre 90 et 99 %. Si l'on compare les séquences isolées directement du génome avec l'élément reconstruit *Aacutmar1*, les homologies sont comprises entre 87 et 90 %. L'élément reconstruit a très probablement été établi à partir de différentes copies de MLE, ce qui peut expliquer des taux d'homologie inférieurs à 90 % avec les "séquences vraies".

Dans les séquences *Aacutmar2* qui ont une taille de 1753 pb, une partie interne de la transposase a été remplacée par un fragment d'ADN n'ayant aucune correspondance avec une séquence connue. La possibilité que cette dernière soit un ET a été envisagée, son analyse a donc été complétée par la détection d'éventuelles répétitions à ses extrémités qu'elle soit orientées dans le même sens ou en sens inverse, mais aucune correspondance n'a été détectée. Dans *Aacutmar2*, il ne s'agit donc pas de l'insertion d'un ET dans un élément tronqué *Aacutmar1*.

L'élément *Aacutmar3* a, quant à lui, perdu une grande partie du gène de la transposase, seul le motif HTH-1 est encore présent dans sa séquence. Il mesure 1006 pb de long mais seule une petite partie des 1110 pb correspond au MLE *Aacutmar*.

## **5.3. Comparaison des séquences nucléotidiques d'*Aacutmar1* à celles de MLE végétaux des banques de données**

Si l'élément *Aacutmar1* présente des similitudes avec les éléments complets isolés chez les plantes et une diatomée (Tab. 14), il est cependant plus petit que les transposons issus des autres espèces. En effet, la longueur moyenne du consensus est de 2114 pb, alors que *Soymar1* chez le soja a une taille de 3491 pb (Jarvik et Lark, 1998), *Osma1* chez le riz de 5259 pb (Tarchiniet al., 2000), *Vulmar1* chez la betterave mesure 3909 pb (Jacobs et al., 2004), *Psmar1* chez le pois de 3204 pb (Macas et al., 2005) et *Phmar1* du bambou 3452 pb (Zhou et al., 2011), *Costa* chez la diatomée a une taille de 3035 pb. Dans ces

séquences si les ITR et les ORF ont des longueurs comparables, on note de grandes différences de taille au niveau des UTR, par exemple l'UTR-5' d'*Aacutimar1* mesure 152 pb mais atteint plus de 1890 pb chez *Osmar1*. Il est difficile d'interpréter les différences de taille au niveau des UTR, toutefois, il serait intéressant de rechercher dans ces séquences d'éventuels promoteurs, par exemple Miskey et al. (2007) ont montré que les 180 pb de l'UTR-5' de l'élément *mariner* isolé de l'humain, *Hsmar1* contient une séquence promotrice.

**Tableau 14 : La longueur des ITR, des UTR, de l'ORF des MLE complets chez les végétaux.**

Élément	Longueur (pb)	ITR-5' (pb)	UTR-5' (pb)	ORF (pb)	UTR-3' (pb)	ITR-3' (pb)
<i>Soymar1</i>	3491	39	952	1275	1186	39
<i>Osmar1</i>	5259	26	1897	1728	1582	26
<i>Vulmar1</i>	3909	32	483	1554	1808	32
<i>Psmar1</i>	3204	31	840	1236	1066	31
<i>Costa</i>	3035	27	1369	1509	106	27
<i>Phmar1</i>	3452	29	490	1983	921	29
<i>ConsAacutmar1</i>	2114	27	152	1473	435	27

### **5.3.1. Le TSD et les ITR des éléments *Aacutmar1***

*Aacutmar1* est bordé par une duplication du site d'insertion (5'-TA-3') qui est une des caractéristiques des MLE. Ses ITR, longs de 27 pb, sont comparables à ceux des MLE complets végétaux disponibles dans les banques de données, dont la taille varie entre 26 pb (élément *Osmar1* du riz *Oryza sativa*) et 39pb (élément *Soymar1* du soja *Glycine max*). Les ITR des différents éléments partagent une homologie élevée au niveau des 16 premiers nucléotides et montrent une séquence bien conservée de cette zone dont le consensus est 5'-CTCCCTCCGT(C/T)CAAAT-3'. Les neuf premiers nucléotides sont eux très conservés entre les différents éléments (*Soymar1*, *Osmar1*, *Vulmar1*, *Psmar1*, *Costa*, *Phmar1*) et dont le consensus est 5'-CTCCCTCCG-3'. Feschotte et al. (2005), ont montré que chez *Oryza sativa*, des mutations localisées parmi les 9 premiers nucléotides (5'-CTCCCTCCG-3') des ITR n'avaient que très peu d'effets sur l'accrochage d'*Osmar 5* sur ses ITR. Ceci indique que ces premières bases des ITR n'ont pas de fonction dans la liaison de la transposase sur ses ITR, cependant leur conservation importante entre les ITR de tous les MLE végétaux alignés suggère tout de même un rôle important de ces nucléotides pour le transposon. Des données similaires sur la conservation importante

d'une partie des ITR ont été mises en évidence dans les MLE isolés chez différentes espèces animales. Ainsi, Lampe et al. (2001) ont effectué une analyse des 28 premières séquences des ITR de 6 sous-familles d'éléments *mariner*. Un logo affichant la fréquence des bases à chaque position a montré que deux régions de l'ITR, situées en positions 3-8 et 14-18, semblent être les plus conservées. La conservation de ces positions suggère que la transposase pourrait établir des contacts spécifiques sur ces régions des ITR. En 2005, Bigot et al., ont réalisé un alignement des ITR d'environ 28-30 pb de nombreux éléments MLE appartenant aux 5 sous-familles (*mauritiana*, *cecropia*, *mellifera/capitata*, *irritans* et *elegans/briggsea*), qui montre le consensus 5'-YYAGRT-3' au niveau de leurs extrémités extérieures, consensus correspondant au signal de coupure de l'élément. Cependant les auteurs indiquent également que lorsque l'alignement est fait entre les membres d'une même sous-famille, les extrémités sont alors bien plus conservées. Il est donc possible que les MLE isolés chez les végétaux et présentant des extrémités des ITR très conservées appartiennent tous à la même sous-famille. Bigot et al. (2005) ont montré également sur les séquences des ITR, la présence de séquences palindromiques et de motifs miroirs. Les séquences palindromiques seraient corrélées à la liaison de la transposase aux ITR et les motifs miroirs aideraient à la courbure des ITR stabilisant les interactions ADN/protéine. Dans le cas des ITR d'*Acutmar* et des autres MLE végétaux, de telles structures n'ont pas été détectées.

Les ITR-5' et 3' des éléments *Acutmar* ne sont pas totalement identiques, ils possèdent quatre différences aux positions 2, 15, 21 et 22, ils sont donc imparfaits. Des ITR imparfaits sont fréquemment mis en évidence dans les éléments de type *mariner* (Plasterk et al., 1999 ; Brillet et al., 2007). Par exemple, *Mos1* qui est un élément autonome isolé de *Drosophila mauritiana* possède des ITR inversés imparfaits de 28 pb avec 4 différences. Dans les MLE complets de végétaux, les éléments *Soymar1* du soja et de la diatomée marine *P. multiseriata* montrent respectivement 2 et 3 différences entre leur ITR-5' et leur ITR-3' donc, tout comme *A. acutiuscula*, les MLE de ces deux espèces possèdent des ITR imparfaits. En revanche, les ITR des éléments *Osmar1* du riz, *Vulmar1* de la betterave, *Phmar1* du bambou et *Psmar1* du pois sont parfaits, c'est-à-dire ne présentant aucune différence entre l'ITR-5' et l'ITR-3'.

### **5.3.2. Structure de la transposase *AaTnp***

#### ***Les motifs du domaine de liaison à l'ADN***

La transposase *AaTnp* d'*A. acutiuscula* possède, comme toute enzyme de ce type, deux domaines : un domaine de liaison à l'ADN en région N-terminale et un domaine catalytique en région C-terminale. L'existence des motifs HTH dans le domaine N-terminal est parfois discutée dans la littérature, les logiciels qui permettent la détection des HTH sont peu adaptés aux ET. Toutefois, Plasterket al.(1999) ont montré que les transposases des TLE (*Tc1*-Like Element) et celles des MLE possèdent deux motifs HTH. Zhanget al.(2001) ont rapporté qu'une partie du motif HTH putatif situé entre les résidus 88 et 108 est essentielle pour l'activité de liaison à l'ADN de la transposase *Mos1*. Avec le programme de prédiction des motifs HTH utilisé pour analyser les transposases des MLE de plantes, dont *Aacutmar1*, il a été constaté que le score de prédiction indique la probabilité maximale (100 %) pour que *AaTnpase* contiennent deux motifs HTH. Dans la majorité des transposases de la superfamille *Tc1-mariner* décrites à ce jour, essentiellement chez des espèces animales, un bloc de quatre acides aminés "GRPR" se trouve localisé au milieu de la zone séparant les deux HTH (Plasterk et al., 1999). Dans les transposases des végétaux dont les diatomées, ce bloc est remplacé par le tripeptide GRK, qui est hautement conservé au centre de la zone séparant les deux HTH. Que ce soit le motif GRPR ou le motif GRK, leur fonction n'est pas connue à ce jour. Toutefois, au vu de leur très forte conservation dans les séquences, ils ont très probablement un rôle primordial dans le domaine de liaison à l'ADN des transposases.

Un second motif situé dans le domaine de liaison à l'ADN est primordial pour le fonctionnement de la protéine, il s'agit du NLS qui doit être présent pour que la transposase, enzyme nucléaire puisse entrer dans le noyau après avoir été produite dans le cytoplasme. La possession d'un NLS indique que les transposons peuvent profiter de la machinerie de transport de cellules hôtes pour l'absorption nucléaire de leurs transposases (Ivics et al., 1996). Dans tous les MLE complets de végétaux, un NLS a pu être prédit à l'aide du logiciel utilisé. En alignant chez les végétaux, les transposases déduites des éléments complets, nous avons pu localiser l'emplacement d'un NLS dans chacune des six transposases. Bien que ces motifs soient assez différents les uns des autres, ils partagent une même caractéristique, ils possèdent quatre acides aminés en commun avec le deuxième motif HTH. Le chevauchement des motifs HTH et NLS est décrit dans la littérature dans de nombreuses transposases de la superfamille *Tc1-mariner* (Plasterk et

al., 1999). L'alignement des MLE complets issus des végétaux a permis de confirmer le chevauchement du deuxième motif HTH et du NLS.

Dans la structure du NLS de l'élément *Mos1*, l'acide aminé R (Arginine) est essentiel pour l'activité de la transposase (Lohe et al., 1997), un acide aminé R au début des NLS des MLE complets des végétaux est également présent et bien conservé à l'exception de l'élément *Costa* qui en est dépourvu. Dans les MLE isolés de végétaux, nous avons observé plusieurs résidus très conservés, L, K, R en position 19, 24, 27, respectivement mais le rôle de chacun d'eux dans la protéine n'est pas identifié.

#### ***Les motifs du domaine catalytique***

Les motifs entourant les trois D sont respectivement "HVDEKWF", "IQQDGSTAH" et "QSPDTNIDLAFF", ils sont donc légèrement différents des motifs canoniques des transposases de végétaux "HIDEKWF", "IQQDNARTH" et "NSPDL/MNLDLGFF" de la même famille.

Les transposases *AaTnp* de l'élément *Aacutmar1* présentent donc des caractéristiques suggérant leur fonctionnalité : plusieurs d'entre elles possèdent des ORF, deux HTH, un NLS ainsi que des motifs conservés dans la partie catalytique qui font d'elles des enzymes potentiellement actives.

La présence d'un motif DD43D dans la séquence de la transposase *AaTnp* au lieu de la triade catalytique DD39D trouvée dans la plupart des MLE chez les plantes terrestres est tout à fait remarquable. Ce motif DD43D est à ce jour unique parmi les éléments mariner connus de la lignée verte. Les implications de cette particularité sont reprises dans le paragraphe suivant qui traite des relations phylogénétiques des séquences *Aacutmar1* avec les transposons de la superfamille *Tc1-mariner*.

#### **5.4. Dénombrement des copies d'*Aacutmar* dans le génome hôte**

Une analyse réalisée ici par PCR quantitative a permis d'estimer le nombre de transposons *Aacutmar* à environ  $21 \pm 8$  copies dans le génome de la diatomée *A. acutiuscula*. Celles-ci ont été dénombrées en utilisant des amorces spécifiques de la région conservée des transposases *mariner*. Il faut encore noter que l'estimation faite ici ne tient pas compte des éléments *Aacutmar2* et 3 car les amorces qui ont permis le dénombrement sont localisées au niveau d'une région de la transposase qui a été perdue

par ces deux types d'éléments. Inversement, l'estimation tient compte des éléments *Aacutmar4* qui sont tronqués à l'extrémité 3' mais ils contiennent bien la zone d'accrochage des amorces. Pour estimer les quatre types d'éléments simultanément, il serait nécessaire de définir des amorces au niveau du motif HTH n°1 qui leur est commun. L'estimation du nombre de copies a été entreprise avant l'obtention des éléments amplifiés à l'aide de ITR, la localisation des amorces a alors été choisie dans la région catalytique de la transposase, zone généralement très conservée sauf, bien-sûr, si elle est délétée, ce qui est le cas pour *Aacutmar2* et 3.

Chez *A. acutiuscula*, la quantification des copies de MLE indique un nombre similaire, environ 21, à ce qui est décrit dans la littérature pour les génomes de plantes. Par exemple, chez le soja, 25 copies de l'élément *Soymar1* ont été observées par Southern-blot (Jarvik et al., 1998). Dans le riz, *Oryza sativa*, 34 MLE dont 22 pleine taille contenant les ITR, le TSD et le gène de la transposase ont été détectées à partir des génomes disponibles dans les banques de données (Feschotte et al., 2003). Dans le bambou *P. pubescens*, Zhou et al. (2010) ont estimé à environ une vingtaine de copies le nombre de MLE par Southern-blot. En utilisant les génomes séquencés, Zhou et al. (2011) ont obtenu des résultats similaires avec une quinzaine d'éléments identifiés. Une étude récente de Han et al. (2013) indique le nombre de copies de la superfamille *Tc1-mariner* dans quatre céréales *Brachypodium distachyon*, le riz (*O. sativa*), le sorgho (*Sorghum bicolor*) et le maïs (*Zea mays*), qui a été estimé à 36, 45, 57 et 47 éléments respectivement. Ainsi, les données recueillies indiquent entre 20 et 50 copies de MLE par génome chez les plantes. Par contre, le nombre de MLE mis en évidence dans le génome de certains animaux et notamment chez certains insectes est beaucoup plus élevé. En effet, on peut citer l'exemple du papillon *Hyalophora cecropia*, dont les MLE ont été estimé à environ 1000 copies par génome (Lidholm et al., 1991), celui du moustique *Ochlerotatus atropalpus* (Zakharin et al., 2004) chez lequel l'utilisation de la technique d'hybridation slot-blot a permis d'estimer le nombre d'éléments *Atmar1* et *Atmar2* respectivement entre 370 à 1200 et de 100 à 300 copies ou encore le nombre de MLE chez la sauterelle *Eyprepocnemis plorans* estimé à  $344 \pm 44$  copies par Q-PCR (Montiel et al., 2012). Toutefois, les MLE ne sont pas toujours si nombreux dans les génomes animaux, en effet leur nombre varie de quelques copies à plusieurs milliers par génome (Capy et al., 1991 ; Lidholm et al., 1991 ; Lohe et al., 1995 ; Robertson et Lampe, 1995). S'il en est de même chez les végétaux, il est possible que certains génomes soient plus



envahis que ceux des plantes citées ci-dessus mais actuellement les données dans la littérature manquent à ce sujet. Nous pouvons rappeler ici qu'aucun élément apparenté à un MLE n'a été mis en évidence dans les quatre génomes séquencés de diatomées disponibles dans les banques de données (*P. tricornutum*, *T. pseudonana*, *Fragilariopsis*, *P. multiseriata*) (Rouault et al. en préparation).

Toutefois, l'estimation du nombre de copies d'un ET n'est pas toujours aisée et ce, quelle que soit la méthode utilisée. L'estimation peut être réalisée par : hybridation d'une sonde marquée (Southern-blot ou slot-blot) permettant de visualiser directement l'envahissement du génome par l'élément recherché, par l'analyse des génomes séquencés disponibles dans les banques de données ou encore par PCR quantitative. Pour le comptage, l'hybridation d'une sonde spécifique marquée est une méthode efficace lorsque le nombre de copies est faible de 1 à 50 environ copies car au-delà l'hybridation produit un "smear" dans lequel il est difficile de distinguer les différentes copies, qui ne sont alors plus comptabilisables. Si le génome contient des éléments différents les uns des autres qu'ils soient intacts, mutés ou tronqués, pour en dénombrer le plus possible, il est intéressant de faire un mélange de sondes permettant d'hybrider les différentes copies.

L'estimation du nombre de copies à partir d'un génome séquencé dépend de la qualité du séquençage et aussi de celle de l'assemblage. Par exemple, les ET peuvent se trouver dans des zones difficiles à assembler et donc être écartés du processus, ils ne seront pas décelables par des analyses *in silico* bien que présents dans le génome séquencé.

L'estimation par PCR quantitative repose sur l'amplification des copies par des amorces spécifiques et peu éloignées l'une de l'autre, si la zone est délétée la copie ne sera pas comptabilisée. Pour palier à ce problème, il est possible de faire des comptages utilisant des jeux d'amorces différents le long de la séquence de l'élément, optimisant ainsi la probabilité d'en amplifier le plus grand nombre. Pour avoir une idée la plus précise possible du nombre de copies de transposons dans un génome, l'idéal est donc de pouvoir combiner les 3 approches quand cela est possible. La donnée qui manque le plus fréquemment est celle du séquençage complet du génome et de son annotation mais actuellement de plus en plus de génomes sont disponibles dans les banques de données.

Chez *A. acutiuscula*, le nombre élevé de MLE présents dans le génome et la forte proportion d'éléments à ORF, parmi les 12 séquences d'*Aacutmar1*, six possèdent un cadre de lecture naturel susceptible de coder une transposase active, suggèrent une mobilité récente ou ayant disparu récemment des éléments *Aacutmar1*. Ces données

peuvent être interprétées comme une phase d'accumulation des MLE dans le génome hôte. En effet, le nombre de copies d'une famille d'ET passe par trois phases dans un génome, phases intégrées dans le cycle de vie de l'ET (Lohe et al., 1995 ; Kidwell et Lisch, 2001). Les trois phases successives sont : 1) l'invasion caractérisée par l'amplification du nombre de copies ; 2) la maturité qui correspond à la mise en place de systèmes de régulation de la transposition avec par exemple, l'apparition de mutations dans les copies limitant ainsi leur mobilité et 3) la sénescence montrant la diminution des éléments par élimination et accumulation. Toutefois, la présence des éléments mutés *Aacutmar2* et 3 peut aussi laisser penser que la phase d'accumulation est atteinte et que le génome est proche de sa maturité. Pour trancher entre les phases, accumulation ou maturité, il serait possible de suivre plusieurs générations d'algues, d'analyser leur contenu en MLE pour savoir si le nombre de copies poursuit son amplification ou non.

## **5.5. Relations phylogénétiques d'*Aacutmar1***

Comprendre l'évolution des ET et leurs relations phylogénétiques est une étape essentielle pour comprendre l'évolution des génomes dont ils sont issus. Notre analyse phylogénétique est concentrée sur l'histoire évolutive des MLE des diatomées par rapport aux MLE d'autres protistes et de végétaux et ce, dans l'ensemble de la superfamille *Tc1-mariner* des eucaryotes.

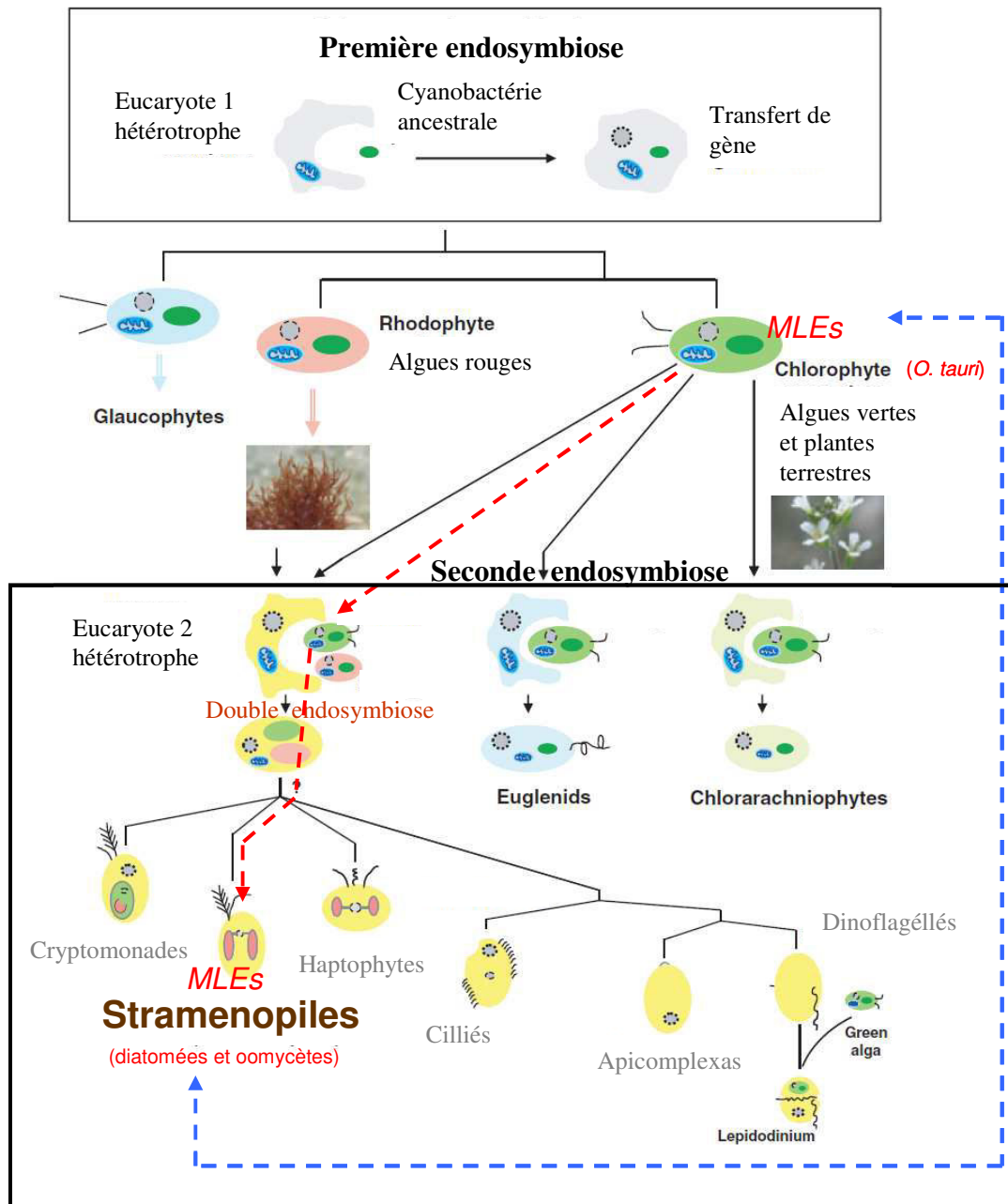
### **5.5.1. Révision de la nomenclature du groupe formé par les MLE de végétaux et de Straménopiles**

Jusqu'à présent, le clade qui rassemblait les séquences de MLE isolés des génomes végétaux était référencé dans la littérature sous la dénomination famille II-DD39D (Shao et Tu, 2001) ou par plant MLE (Feschotte et Wessler, 2002) ou encore *ITmD39D* (Brillet et al., 2008). L'analyse phylogénétique réalisée ici avec toutes les séquences de transposases complètes de Chlorophytes (algues vertes), de plantes et de Straménopiles (diatomées et oomycètes) est la première analyse exhaustive des transposases MLE d'organismes photosynthétiques. En effet, les précédentes analyses étaient focalisées sur les plantes et ne concernaient que des portions de la transposase (Feschotte et Wessler 2002, Jacobs et al. 2004, Zhou et al. 2010, 2011). Notre analyse montre que les séquences de transposases MLE issues d'organismes photosynthétiques, bien que classées dans la superfamille *Tc1-mariner*, présentent diverses structures de triades catalytiques, DD39D,

DD43D, ou encore DD45D, et ne peuvent donc plus être regroupées sous le terme DD39D. Il apparaît clairement qu'une nouvelle dénomination plus appropriée à l'ensemble des organismes photosynthétiques devra être utilisée et qu'elle ne pourra être basée sur le nombre d'acides aminés situés entre les deux derniers résidus aspartique de la triade. Ceci a déjà été évoqué par Rouault et al. (2012) lors de travaux sur la classification des éléments de la superfamille *Tc1-mariner*, ces auteurs suggèrent que les MLE de Straménopiles et de plantes pourraient être regroupés sous une dénomination de famille plus générale. Le terme de famille *Chlorophyllis* sera probablement proposé car il permet d'évoquer l'origine chlorophyllienne des espèces hôtes. Le présent travail contribuera à une publication de Rouault et al. actuellement en préparation et concernant la révision de la classification de la superfamille *Tc1-mariner*.

#### **5.5.2. Relations phylogénétiques entre les MLE des diatomées et les MLE dans la lignée verte**

Au vu des nombreux motifs communs, il n'est pas surprenant que les MLE d'*A. acutiuscula* se regroupent avec les MLE de plantes. En effet, les MLE de diatomées sont regroupés dans un clade qui contient les MLE de plantes et de Straménopiles (oomycètes et diatomée séquencée *P. multiseriis*). Ce résultat est toutefois inattendu dans le sens où les hôtes sont issus de deux lignées phylogénétiquement éloignées : celle issue d'une endosymbiose primaire (Falkowsky et al., 2004 ; Reyes-Prieto et al., 2007) et celle des Straménopiles regroupant les organismes issus d'une endosymbiose secondaire d'une algue rouge et peut être d'une algue verte (Falkowsky et al., 2004 ; Kutschera et Nikklas, 2005 ; Moustafa et al., 2009 ; Cavalier-Smith, 2010 ; Tirichine et Bowler, 2011 ; Deschamps et Moreira, 2012). Deux hypothèses peuvent expliquer cette proximité phylogénétique, la première est la transmission verticale des MLE dans la lignée des Straménopiles et la seconde un transfert horizontal entre la lignée des plantes et celle des Straménopiles. Ces deux hypothèses sont détaillées ci-dessous en s'appuyant sur la figure 71.



**Figure 71 : Schéma récapitulatif des deux hypothèses pouvant expliquer la transmission des MLE entre les algues vertes et les Straménopiles (d'après la figure de Tirichine et Bowler, 2011).**

Deux hypothèses sont proposées : la première est un transfert vertical (indiquée par les flèches en pointillés rouges) depuis une algue verte (Chlorophyte, Prasinophyceae) ancestrale de la lignée des plantes en se basant sur la théorie de la double endosymbiose secondaire à l'origine des Straménopiles. La deuxième hypothèse est représentée par la flèche en pointillés bleus et correspond à un transfert horizontal des MLE entre les algues vertes vers les Straménopiles.

La première hypothèse pour expliquer la proximité phylogénétique de ces éléments concerne la transmission par transfert vertical des MLE qui auraient été présents très tôt dans la lignée des eucaryotes. En tenant compte des récentes hypothèses sur la possible double endosymbiose secondaire d'une algue rouge (Rhodophyte) et d'une algue verte (Prasinophyceae) ayant donné lieu à la lignée brune (Becker et al., 2008 ; Moustafa et al., 2009 ; Tirichine et Bowler, 2011), nous pouvons supposer que les MLE étaient présents chez l'algue verte ayant participé à cette double endosymbiose. En effet, les MLE ont été mis en évidence chez une algue verte *O. tauri* (Prasinophyceae) mais pas dans les deux génomes de Rhodophytes actuellement disponibles dans les banques de données (Matsuzaki et al., 2004 ; Nozaki et al., 2007 ; Nakamura et al., 2013). Ainsi, nous pouvons supposer que les MLE ont été transmis par l'algue verte au moment de l'endosymbiose secondaire dans la lignée brune. Comme d'autres gènes, ils auraient alors été transférés dans le noyau de l'hôte ancestral, expliquant leur présence chez ses descendants notamment chez les diatomées et les oomycètes phylogénétiquement proches de ces dernières (Harper et al., 2005). Toutefois, la présence des MLE n'a pas été détectée chez toutes les espèces dérivant de l'endosymbiose secondaire et composant actuellement la lignée brune. Lorsque des données sont disponibles dans la littérature ou dans les banques de données, nous avons pu noter que ces transposons n'ont été détectés que ponctuellement chez les algues vertes et chez les plantes comme le montre l'échantillon de notre analyse. Par exemple un seul MLE est connu à l'heure actuelle chez les Chlorophytes avec *O. tauri* et ils n'ont pas été détectés dans le génome d'autres membres de ce groupe tels que *Chlamydomonas reinhardtii*, *Volvox carteri* ou *Chlorella vulgaris*. Rappelons que chez *C. reinhardtii*, d'autres ET ont été mis en évidence (Dai et al., 1988 ; Ferris, 1989 ; Graham et al., 1995 ; Wang et al., 1998). De la même façon, les MLE n'ont pas été détectés dans d'autres génomes issus d'une endosymbiose secondaire hors Straménopiles. Cette distribution sporadique des MLE dans ces deux lignées peut être expliquée par la perte des éléments au cours de l'évolution (Capy et al., 1994 ; Lohe et al., 1995 ; Pace et al., 2008).

La distribution disparate des MLE dans les organismes issus soit d'une endosymbiose primaire soit d'une endosymbiose secondaire nous amène à supposer que les MLE ont également pu, selon une deuxième hypothèse, être transférés horizontalement entre une algue verte et l'ancêtre des Straménopiles. Ce type de transfert aurait pu avoir lieu entre une Prasinophyceae (comme *O. tauri*) et l'ancêtre de la lignée des Straménopiles.

Comme indiqué par Fortune et al. (2008), le transfert horizontal peut être identifié par une incongruité entre la phylogénie des ET et la phylogénie des hôtes, ce qui est le cas des MLE étudiés ici. Bien que ce mode de transfert entre eucaryotes soit largement décrit, il est bien décrit chez les procaryotes. Néanmoins, le transfert de MLE entre drosophiles semble plausible (Picot et al., 2008 ; Rubin et al., 2011 ; Wallau et al., 2011 ; Wallau et al., 2012). De plus, il a été suggéré par certaines analyses que des gènes de diatomées et de oomycètes ont été acquis via ce mode de transfert aux moyens de vecteurs tels que les virus (Bowler et al., 2008 ; Sobecky et Hazen, 2009) ou bien par des vecteurs qui pourraient être les plasmides retrouvés dans le milieu marin et dans les diatomées marines (Hildebrand et al., 1991). Nous pouvons aussi émettre l'hypothèse que le transfert horizontal de gènes de résistance au froid entre basidiomycètes (Opisthocontes) et la diatomée *F. cylindrus* (Sorhannus, 2011) pourrait aussi concerner des MLE. Il est à noter également que de récents travaux montrent que la distribution des ET dans divers génomes de plantes serait le résultat de transferts horizontaux fréquents, plusieurs exemples ont été apportés par l'étude des ET dans les génomes de riz (Roulin et al., 2009 ; Schaack et al., 2010). Ainsi, il apparaît que cette hypothèse d'un transfert horizontal des MLE ne peut être écartée et pourrait expliquer la proximité phylogénétique des MLE de la diatomée *A. acutiuscula* et de Straménopiles en général avec les MLE identifiés dans des génomes de plantes comme le montre notre analyse phylogénétique. Pour trancher cette question, il faudrait pouvoir dater les différents événements qui ont ponctué l'histoire évolutive des MLE.

Il apparaît évident qu'il y a une pénurie de données concernant la présence des MLE dans les diverses lignées et que leur détection dans plusieurs génomes séquencés notamment d'algues rouges et de Glaucophytes permettrait de préciser les relations phylogénétiques présentées dans notre analyse. Les relations phylogénétiques des MLE fournissent un aperçu complémentaire de l'évolution des organismes photosynthétiques. En effet, les résultats obtenus dans cette étude ne sont pas en contradiction avec l'hypothèse de la double endosymbiose secondaire qui ne fait pas l'unanimité dans la communauté scientifique (Deschamps et Moreira, 2012). Il est nécessaire de continuer à rechercher et à analyser les relations phylogénétiques d'autres MLE issus de génomes complets de plusieurs lignées pour mieux comprendre leur évolution chez les eucaryotes photosynthétiques.

## **Chapitre 2 : Impact de variations de température et d'excès de métaux sur l'expression des MLE**

### **1. Impact des traitements thermiques**

#### **1.1. Activité photosynthétique**

##### **1.1.1. Intensité photosynthétique**

Après différentes durées de traitement thermique (0 ; 0,5 ; 2, 5 et 8 h), l'intensité photosynthétique (IP) des échantillons a été déterminée par la mesure du dioxygène dégagé. Toutes les mesures ont été répétées au moins trois fois avec des cultures différentes. Les résultats sont présentés dans le tableau 15 en fonction de la durée du traitement et des températures appliquées (16, 24 et 32 °C), 24 °C étant la température de référence.

**Tableau 15 : Effet de la température sur l'intensité photosynthétique ( $\mu\text{mol O}_2 \text{ h}^{-1} \text{ mg}^{-1} \text{ Chl } a$ ) de la diatomée *A. acutiuscula* maintenue pendant 0 à 8 h à 16, 24 et 32 °C.**

Moyennes  $\pm$  SE (n= 3). Chaque condition thermique (16 et 32 °C) a été comparée avec le témoin à 24 °C, les valeurs non significativement différentes (test TukeyHSD,  $p_{\text{adj}} \leq 0,05$ ) partagent la même lettre.

Température	Durée du traitement thermique				
	0h	0.5h	2h	5h	8h
16 °C	297 $\pm$ 15	<b>78 <math>\pm</math> 4</b>	95 $\pm$ 5	<b>67 <math>\pm</math> 10</b>	<b>58 <math>\pm</math> 5</b>
24 °C	264 $\pm$ 10	260 $\pm$ 5	296 $\pm$ 5	240 $\pm$ 8	276 $\pm$ 20
32 °C	297 $\pm$ 15	332 $\pm$ 24	268 $\pm$ 20	293 $\pm$ 36	172 $\pm$ 8

Les valeurs de l'IP des microalgues en condition témoin (24 °C) ne varient pas au cours de l'expérimentation (Tab. 15). Chez les diatomées qui étaient cultivées à 24 °C puis placées à 16 °C, la diminution de 73 % de l'IP est très nette et significativement différente de la valeur témoin dès 30 min à 16 °C, puis l'IP continue de diminuer jusqu'à la fin du traitement où elle ne représente plus que 20 % de l'IP du témoin. Cet abaissement de température affecte donc le métabolisme de cette espèce.

Contrairement à ce qui est observé à 16 °C, le dégagement d'O<sub>2</sub> à 32°C n'est pas significativement différent de celui mesuré en condition témoin (24 °C).

### 1.1.2. Fluorescence de la chlorophylle *a*

Parallèlement aux mesures d'IP, les paramètres de la fluorescence de la Chl *a* des différents échantillons ont été mesurés. A la fin des différents traitements thermiques, un échantillon de 3 mL de culture homogénéisée a été mis à l'obscurité dans une cuve bi-miroitée pendant 15 min afin de réoxyder tous les accepteurs d'électrons du PSII, tous les centres réactionnels des photosystèmes sont alors ouverts. Puis la fluorescence minimale à l'obscurité ( $F_o$ ) et maximale à l'obscurité ( $F_m$ ) ont été mesurées, les rendements quantiques maximum ( $F_v/F_m$ ) ont été calculés (Tab. 16). Toutes les mesures ont été répétées trois fois avec des cultures différentes et comparées aux paramètres obtenus avec les cellules maintenues à la température témoin 24 °C dans chaque cas.

**Tableau 16 : Effet de la température sur le rendement quantique maximum ( $F_v/F_m$ ) chez *A. acutiuscula* maintenue pendant 0 à 8 h à 16, 24 et 32 °C.**

Moyennes  $\pm$  SE (n= 3). Chaque condition thermique (16 et 32 °C) est comparée séparément avec le témoin à 24 °C, les valeurs non significativement différentes (test Tukey HSD,  $p_{adj} \leq 0,05$ ) partagent la même lettre.

Température	Durée du traitement thermique				
	0h	0.5h	2h	5h	8h
16 °C	0,600 <sup>a</sup> $\pm$ 0,004	0,572 <sup>a</sup> $\pm$ 0,008	0,595 <sup>a</sup> $\pm$ 0,019	0,501 <sup>a</sup> $\pm$ 0,022	0,436 <sup>a</sup> $\pm$ 0,008
24 °C	0,596 <sup>a</sup> $\pm$ 0,011	0,570 <sup>a</sup> $\pm$ 0,026	0,584 <sup>a</sup> $\pm$ 0,010	0,593 <sup>a</sup> $\pm$ 0,026	0,593 <sup>a</sup> $\pm$ 0,014
32 °C	0,597 <sup>a</sup> $\pm$ 0,006	0,597 <sup>a</sup> $\pm$ 0,01	0,605 <sup>a</sup> $\pm$ 0,005	0,563 <sup>a</sup> $\pm$ 0,005	0,558 <sup>a</sup> $\pm$ 0,001

Après l'obtention des paramètres  $F_v/F_m$ , les échantillons ont été exposés à la lumière actinique (300  $\mu$ mol photons  $m^{-2} s^{-1}$ ) pendant 20 min. La fluorescence stabilisée ( $F_s$ ) et maximale à la lumière ( $F_m'$ ) sont mesurées et permettent de déterminer le rendement quantique effectif  $\Phi_{PSII}$  (Tab. 17).

**Tableau 17 : Effet de la température sur le rendement quantique effectif ( $\Phi_{PSII}$ ) chez *A. acutiuscula* maintenue pendant 0 à 8 h à 16, 24 et 32 °C.**

Moyennes  $\pm$  SE (n= 3). Chaque condition thermique (16 et 32 °C) est comparée séparément avec le témoin à 24 °C, les valeurs non significativement différentes (test Tukey HSD,  $p_{adj} \leq 0,05$ ) partagent la même lettre.

Température	Durée du traitement thermique				
	0h	0.5h	2h	5h	8h
16 °C	0,222 <sup>a</sup> $\pm$ 0,003	0,155 <sup>a</sup> $\pm$ 0,007	0,155 <sup>a</sup> $\pm$ 0,007	0,193 <sup>a</sup> $\pm$ 0,007	0,183 <sup>a</sup> $\pm$ 0,010
24 °C	0,234 <sup>a</sup> $\pm$ 0,016	0,260 <sup>a</sup> $\pm$ 0,004	0,264 <sup>a</sup> $\pm$ 0,008	0,256 <sup>a</sup> $\pm$ 0,004	0,267 <sup>a</sup> $\pm$ 0,008
32 °C	0,222 <sup>a</sup> $\pm$ 0,003	<b>0,292<sup>b</sup><math>\pm</math> 0,011</b>	0,219 <sup>a</sup> $\pm$ 0,004	0,202 <sup>a</sup> $\pm$ 0,003	0,202 <sup>a</sup> $\pm$ 0,003



Les paramètres  $F_v/F_m$  et  $\Phi_{PSII}$  des témoins à 24 °C ne varient pas significativement au cours de l'expérimentation. De même,  $F_v/F_m$  n'est pas significativement affecté par les variations de température appliquées. Pour les échantillons exposés à 16 °C, nous observons une réduction des rendements quantiques, dès 30 min de traitement, mais qui n'est pas significative. Un traitement à 32 °C induit une augmentation très marquée (+ 31 %) et significative de  $\Phi_{PSII}$  après 30 min à cette température, puis ce paramètre se stabilise à une valeur non significativement différente de celle des témoins jusqu'à la fin de l'expérimentation.

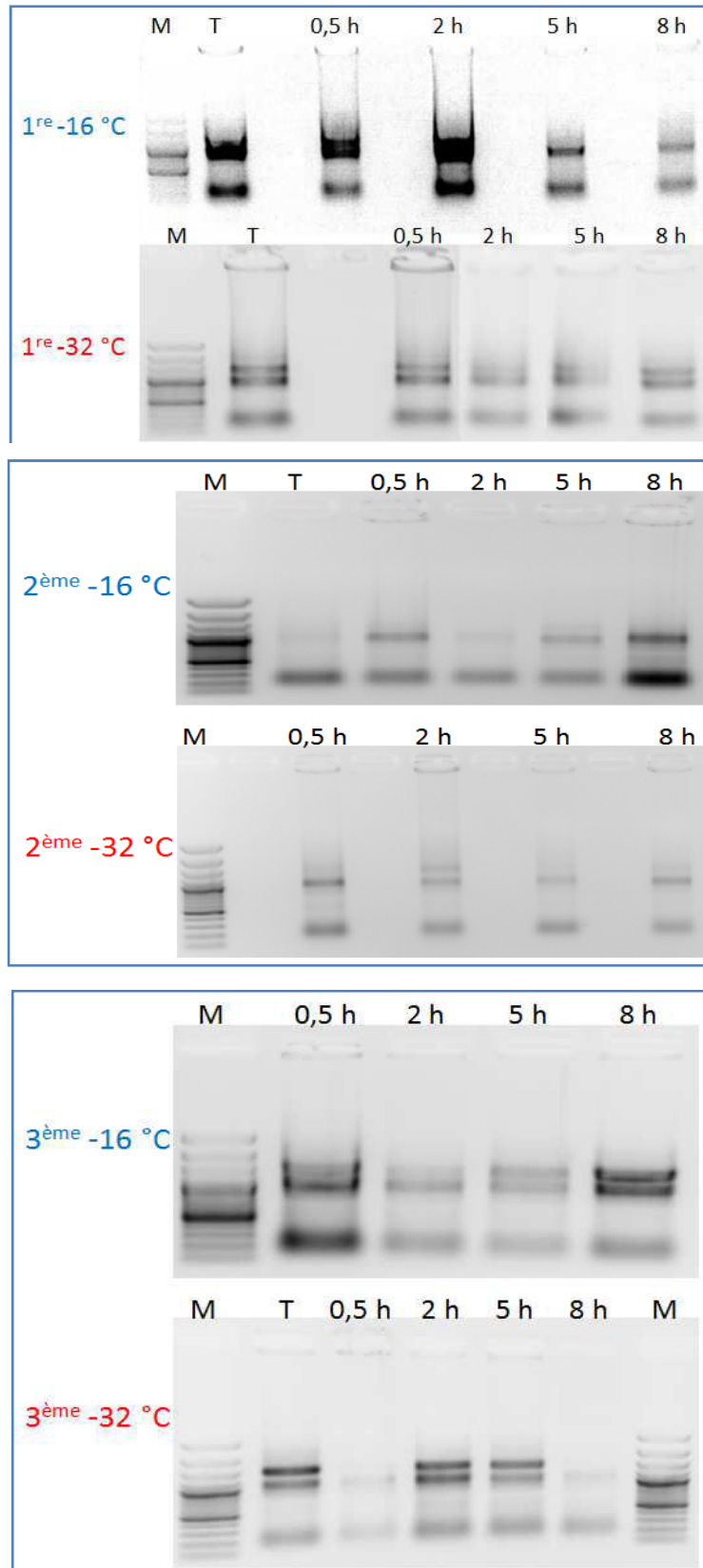
Les mesures d'intensité photosynthétique montrent que les microalgues subissent un stress à 16 °C alors qu'elles s'acclimatent à l'augmentation de température à 32 °C.

## **1.2. Détection de l'expression des MLE par RT-PCR**

L'étude physiologique a montré une réponse différente des diatomées aux différentes températures appliquées, nous avons poursuivi ce travail par l'étude de la réponse moléculaire. Pour cela, nous avons caractérisé l'expression de gènes de réponse aux stress (MnSOD et hsp70) et des gènes codant les MLE.

### **1.2.1. Extraction des ARN totaux**

Dans un premier temps, les ARN totaux ont été extraits au TriReagent et leur qualité vérifiée sur gel d'agarose à 1 % (Fig. 72). Pour chaque température (16 et 32 °C) et chacun des trois réplicats, la qualité des ARN a été jugée correcte puisque deux bandes étaient bien visibles et correspondent au 28S (bande supérieure) et 18S (bande inférieure) ribosomiques. Toutefois, le réplicat n° 2 était de moins bonne qualité et présentait une dégradation du 28S. Tous les ARN totaux ont été traités à la DNase puis rétrotranscrits en ADNc avant de servir de matrice pour la PCR.

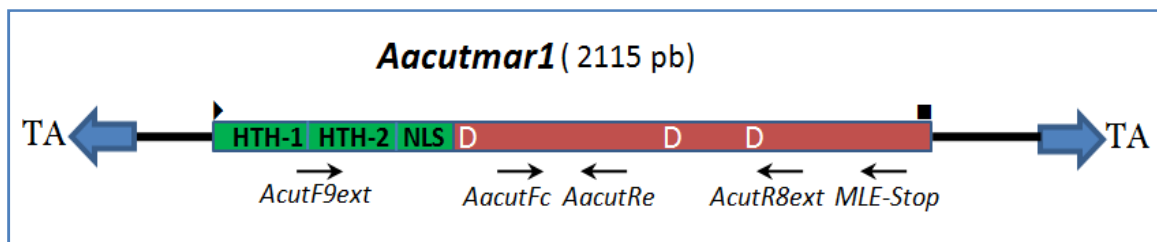


**Figure 72 : Électrophorèse sur gel d'agarose 1 % des produits d'extraction d'ARN totaux d'*A. acutiuscula* placée pendant 0 à 8 h à 16 ou 32 °C. Sont présentés ici 3 réplicats de la cinétique : T= Température témoin 24 °C.**

### 1.2.2. RT-PCR

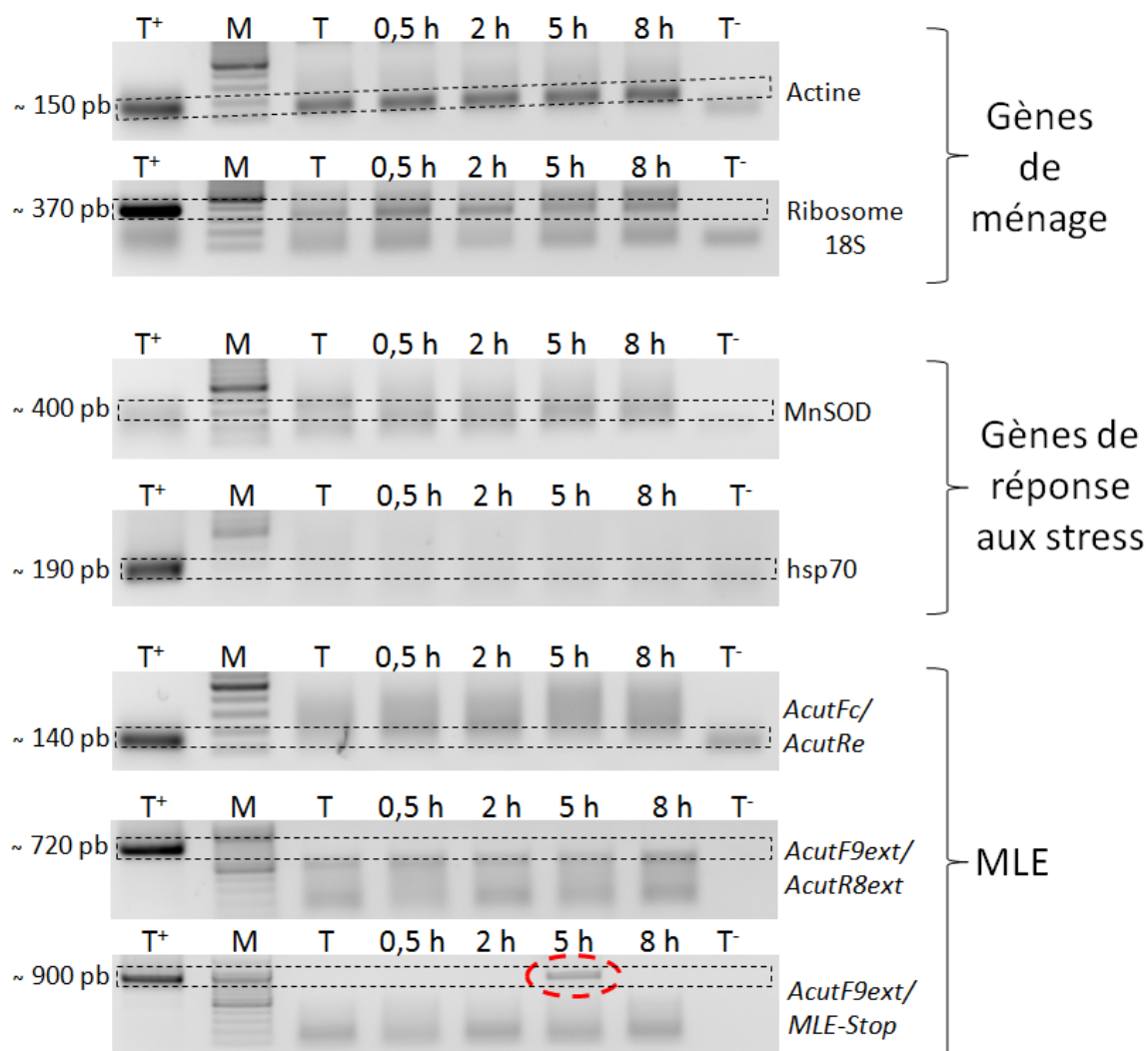
Chaque échantillon ADNc a été amplifié avec un panel de couples d'amorces ciblant différents types de gènes. Les gènes de ménage, actine et ribosome permettent de contrôler la qualité de la rétrotranscription. Les deux gènes de réponse aux stress MnSOD et hsp70 ont permis, d'une part, de s'assurer qu'il existait bien une réponse moléculaire aux conditions thermiques appliquées et d'autre part, pour la MnSOD, qu'aucun ADN génomique contaminant n'était encore présent dans l'extrait d'ARN. Les amorces employées pour amplifier la MnSOD ayant été désignées de part et d'autre d'un intron, absent dans l'ADNc, un produit de PCR montrant une taille supérieure à ce dernier aurait permis de visualiser un ADNg contaminant.

Enfin, en ce qui concerne les MLE, trois couples d'amorces ont été utilisées. La figure 73 présente la localisation de ces amorces sur le gène pleine longueur de l'élément *Aacutmar1*.



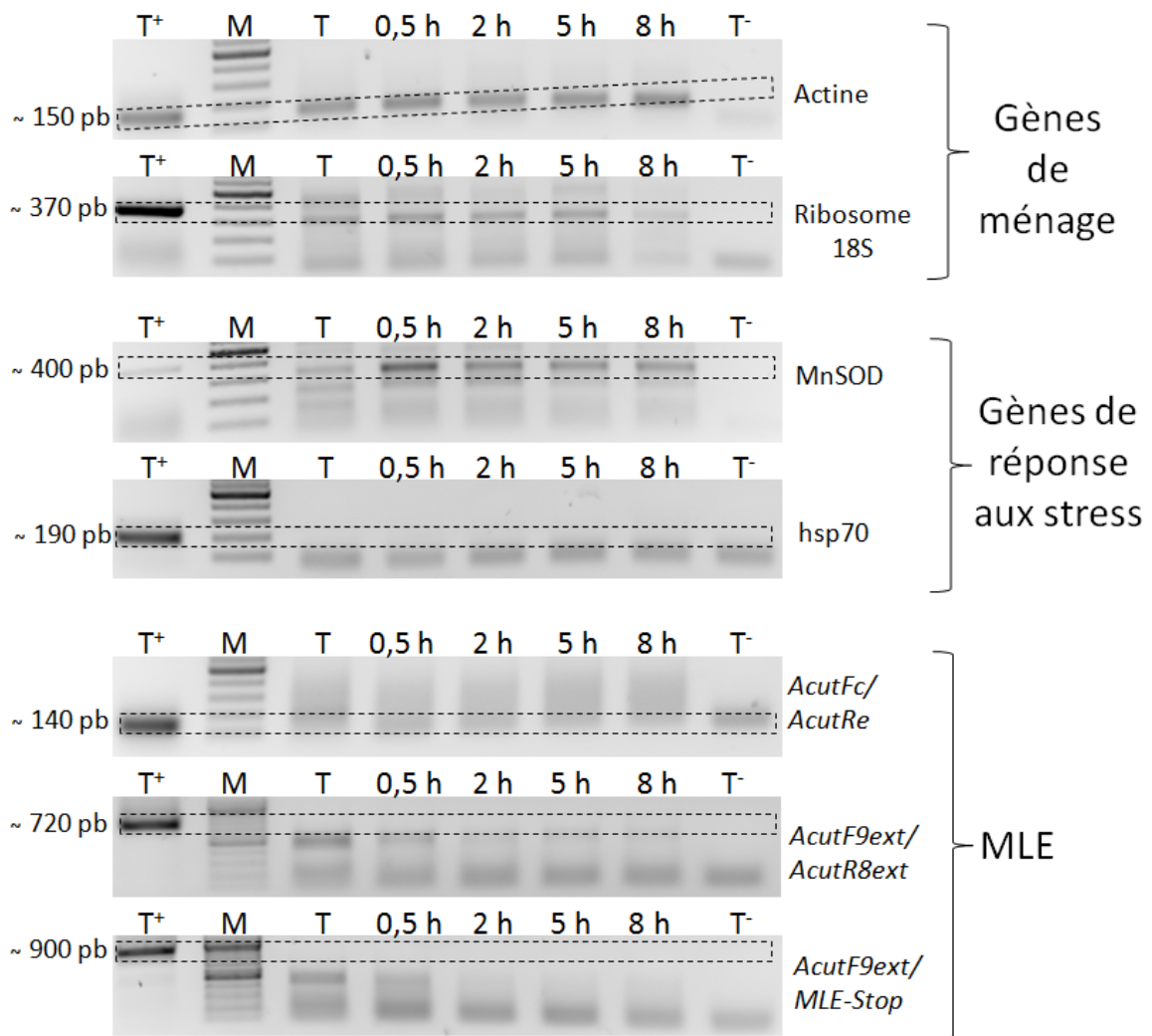
**Figure 73 : Schéma montrant la position des trois couples d'amorces ayant servi à la détection de l'expression des MLE dans les différentes conditions thermiques.**

Pour les gènes de ménage et de réponse aux stress, des amplifications à la taille attendue ont été obtenues aussi bien dans la cinétique à 16 °C que celle à 32 °C (Fig. 74 et 75). Toutefois, quelques exceptions sont à noter comme par exemple pour les hsp70, quel que soit le réplicat en ADNc (Fig. 75) où les amplifications ne sont pas visibles. Sinon les amplifications obtenues étaient de bonne qualité par leur intensité et leur taille, si ce n'est l'amplification du 18S à 8 h et 32 °C pour tous les réplicats qui présente une intensité moindre (Fig. 75).



**Figure 74 : Amplification des gènes de contrôle, de réponse aux stress de 16 °C et des MLE par RT-PCR.**(Exemple de résultats pour le stress 16 °C, 1<sup>er</sup> réplicat).

T<sup>+</sup> : amplification sur ADNg ; M: marqueur de taille 100 bp (Proméga); T : amplification sur ADNc issus de cultures à 24 °C (Témoin); Cinétique de stress (0,5 - 2 - 5 et 8 h de stress); T<sup>-</sup> : témoin négatif sans matrice.



**Figure 75 : Amplification des gènes de ménage, de réponse aux stress de 32 °C et des MLE par RT-PCR.**(Exemple de résultats obtenus à 32 °C, 1<sup>er</sup> réplicat)

T<sup>+</sup> : amplification sur ADNg ; M: marqueur de taille 100 bp (Proméga); T : amplification sur ADNc issus de cultures à 24 °C (Témoin); Cinétique de stress (0,5 - 2 - 5 et 8 h de stress); T<sup>-</sup> : témoin négatif sans matrice

Les produits amplifiés ont été extraits, clonés puis séquencés. Les séquences obtenues pour les ribosomes 18S montraient 100 % d'identité avec les séquences enregistrées dans les banques de données de 18S d'*Amphora sp.*JQ885983 et JF834543. Les séquences d'actine présentaient, quant elles, des pourcentages d'identité de 86 % avec l'actine de *Phaeodactylum tricornutum* (AY729845), ou encore 85 % avec celle de *Nitzschia closterium* (EF553460). L'ensemble des résultats obtenus confirme l'identification des gènes de ménage utilisés.

Pour les gènes de réponse aux stress, MnSOD et hsp70, les résultats obtenus ont été plus délicats à interpréter. En effet, si une amplification a toujours été obtenue avec la matrice ADNg, l'utilisation de la matrice ADNc a donné des résultats plus aléatoires. Pour le gène de la MnSOD, bien que les amplifications en ADNg et ADNc soient de bonne qualité sur le gel concernant les échantillons à 32 °C, les analyses issues du séquençage n'ont pas permis d'identifier ce gène. De même, seules les amplifications avec l'ADNg ont permis d'identifier le gène hsp70. Il se peut cependant que les amplifications ne soient pas visibles sur le gel en ADNc empêchant alors le clonage. Les séquences obtenues pour hsp70 avec l'ADNg pour les deux températures présentaient entre 97 et 100 % d'identité entre elles. Leur identification a été confirmée par l'interrogation des banques de données et a révélé des identités en tblastx avec les hsp70 de *Phaeodactylum tricornutum* (XM\_002177261) et *Thalassiosira pseudonana* (CP001160).

En ce qui concerne les MLE, avec l'ADNc, un seul couple d'amorces (*AcutF9ext/MLE-stop*; Fig. 73) a permis d'obtenir une amplification avec l'échantillon 5 h à 16 °C (Fig. 74). Cinq séquences (*F9-Stop-1* à -5) d'environ 900 pb chacune et présentant entre 81 et 100 % d'identité entre elles ont été obtenues après séquençage. L'interrogation des banques de données a confirmé qu'il s'agissait bien de gènes de transposon de MLE d'*Amphoraacutiuscula* et indiquait une identité de 94 à 95 % avec les clones précédemment isolés *Aacutimar 1.5* (FN546189.1), *1.4* (FN546188.1) et *1.6* (FN546190.1) (Thèse Dorothée Hermann 2011). Le clone *F9-Stop-1* montre de 92 à 93 % d'identité avec les trois clones *Aacutimar*, alors que le clone 3 présente un pourcentage d'identité plus faible d'environ 73 %. L'alignement des 5 séquences ADNc de 900 pb et des 6 séquences pleine taille à cadre de lecture ouvert, identifiées au cours de cette thèse et indiquées dans le chapitre 1, est représenté en figure 76. Les séquences ADNc et pleine taille présentent des identités entre elles comprises de 79 à 99 %. Les clones *F9-Stop-1* et *Aacutimar1.1* et *1.3* sont les plus proches avec 99 et 99,11 % d'identité, confirmant l'identification de ces fragments ADNc.

```

*               480               *               510               *               540               *               570               *               600
F9-Stop-2      : -----ATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTGCTGAAAGAGTTCCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACA : 98
Aacutmar1.12  : ATGAAGAGGAGATCAAATCAAAGAAACATGAGAGAGGAAGAAAGATGGTGTGGGATATTCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTTACTCAGCAAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 608
Aacutmar1.7   : ATGAAGCTGAGATCAAATCAAAGAAACATGAAAGAGGAAGAAAGATGCTGTGGGATATTCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 608
F9-Stop-1     : -----ATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 98
Aacutmar1.2   : ATGAAGCTGAGATCAAATCAAAGAAACATGAAAGAGGAAGAAAGATGCTGTGGGATATTCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 608
Aacutmar1.9   : ATGAAGCTGAGATCAAATCAAAGAAACATGAGAGAGGAAGAAAGATGCTGTGGGATATTCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 608
Aacutmar1.1   : ATGAAGCTGAGATCAAATCAAAGAAACATGAAAGAGGAAGAAAGATGCTGTGGGATATTCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 608
Aacutmar1.3   : ATGAAGCTGAGATCAAATCAAAGAAACATGAAAGAGGAAGAAAGATGCTGTGGGATATTCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTCGCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 608
F9-Stop-3     : -----ATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTTCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACC : 98
F9-Stop-4     : -----ATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTTCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACA : 98
F9-Stop-5     : -----ATATCGATGCAATGGATGAAGCTGTGCAACAAGTTCCTTCTCAGTGAAGAAGCTCTTACCATGGTTTGGCAGCACAGATAGGAGTTCCTCAAGACAACA : 98

*               630               *               660               *               690               *               720               *               750
F9-Stop-2      : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAG---TCGGAAGTCTGTACC : 247
Aacutmar1.12  : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 760
Aacutmar1.7   : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 760
F9-Stop-1     : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 250
Aacutmar1.2   : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 760
Aacutmar1.9   : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 760
Aacutmar1.1   : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 760
Aacutmar1.3   : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGCAAAATAGACCCGAATACGGTAAGGAATGGAAGTTGTCCTCAAGTACTTACC : 760
F9-Stop-3     : CTGTTGCAATACAGAAAGAGGTTCTTTGTTAGACATATCAATGTCCTGAAGCCGACACTCACTCCTGAACACATAGTAGCAAGACTGGAGTTTACCATGAGTAAAGTAGACCCGAATAGGTAAGGAATGGAAG---TCGGAAGTACTTACC : 247
F9-Stop-4     : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGTAAAGTAGACCCGAATACGGTAAGGAATGGAAG---CCCGAAGTACTTACC : 247
F9-Stop-5     : CTGTTGCGATACAGAAAGGAGTTCTTTGTTAGACACAGCAATGCCCTCAAGCCGACACTCACTCCTGAACACAAAGAGCAGCAAGACTTGAGTTTGCCATGAGTAAAGTAGACCCGAATACGGTAAGGAATGGAAG---CCCGAAGTACTTACC : 247

*               780               *               810               *               840               *               870               *               900
F9-Stop-2      : AATGTTAGATGAGGTTTCATGTTGATGAAAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 399
Aacutmar1.12  : AATGTTAGATGAGGTTTCATGTTGATGAAAAGTGTGTTCCACATGACCAAGGATAACCTCGGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 912
Aacutmar1.7   : AATGTTAGATGAGGTTTCATGTTGACGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 912
F9-Stop-1     : AATGTTAGATGAGGTTTCATGTTGACGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 402
Aacutmar1.2   : AATGTTAGATGAGGTTTCATGTTGACGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 912
Aacutmar1.9   : AATGTTAGATGAGGTTTCATGTTGACGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 912
Aacutmar1.1   : AATGTTAGATGAGGTTTCATGTTGACGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 912
Aacutmar1.3   : AATGTTAGATGAGGTTTCATGTTGACGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTCTTCATCTTTGTCAAAGAATGAGGAACCGCCCAAGAAAGAACCTGTCCACACAAATGTTTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGTC : 912
F9-Stop-3     : AATGTTAGATGAGGTTTCATGTTGATGAGAAGTGTGTTCCACATGACAAAGGACAATCAAGTATTCATCTTTGTCAAAGAGCGAGAAACCGCCCAAGAAAGAACCTGTCCACACAAAGGATTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGAA : 398
F9-Stop-4     : AATGTTAGATGAGGTTTCATGTTGATGAAAAGTGTGTTCCACATGACCAAGGATAATCAAGTGTTCATCTTTGTCAAAGAATGAGGAGCCTCCAAAAAGCACTGTTCACACAAAGGATTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGAA : 399
F9-Stop-5     : AATGTTAGATGAGGTTTCATGTTGATGAAAAGTGTGTTCCACATGACCAAGGATAATCAAGTGTTCATCTTTGTCAAAGAATGAGGAGCCTCCAAAAAGCACTGTTCACACAAAGGATTCATTGGCAAAGTGATGTTCTTTGTGTGCACAAGGAA : 399

*               930               *               960               *               990               *               1020               *               1050
F9-Stop-2      : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGCCCTTTGGATGATGTTGGAGAAGGCCAAGGACAATCAAAACACAGACCCAGAGGAGGACCCGGTTTGGAAAGATGAAAGCGTTTAAACAAAGAGGTGATTGCCGA : 551
Aacutmar1.12  : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATTTGGCCCTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCCAAACACAGAGCCAGAGGAGGACCCAGTTTGGAGAAGACGAAAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 1064
Aacutmar1.7   : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATATGTCCTTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCAAAACACAGAGCCAGAGGAGGACCCGGTTTGGAGAGATGAGAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 1064
F9-Stop-1     : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATATGTCCTTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCAAAACACAGAGCCAGAGGAGGACCCGGTTTGGAGAGATGAGAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 554
Aacutmar1.2   : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATATGTCCTTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCAAAACACAGAGCCAGAGGAGGACCCGGTTTGGAGAAGATGAGAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 1064
Aacutmar1.9   : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATATGTCCTTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCAAAACACAGAGCCAGAGGAGGACCCGGTTTGGAGAGATGAGAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 1064
Aacutmar1.1   : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATATGTCCTTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCAAAACACAGAGCCAGAGGAGGACCCGGTTTGGAGAGATGAGAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 1064
Aacutmar1.3   : GTCCAAGGATGGTAGATGGCCGAATGTGGGATGGAAAGATTGGGATATGTCCTTTTGGATGATGTTGGAGAAGGCCAAGAGACGATCAAAACACAGAGCCAGAGGAGGACCCGGTTTGGAGAAGATGAGAGCGTTTAAACAAAGAGGTGTACCCGCCGA : 1064
F9-Stop-3     : GGCCAAGCAT-----GGTGAATGTGGGATGGCAAGCTTGGCCCTTTGGATGATGTTAGGACAGGACAGAGAAATCAAGTAGAGACCGAGCTGGAGGCCCTGTTTGGGAAAGACAGAGTGTGATAGAGACTACTTAAAGACAA : 542
F9-Stop-4     : GGCCGAGGATGGTAATGGCCGAATGTGGGATGGCAAGCTTGGATTTTGGCCCTTTGGATGATGTTGGACAGGACAGAGAGAAATCCGAATACAGAGCCGGCAGGAGGCCCTGTTTGGGAAAGATGAGAGTGTGATAGAGACACTTAAAGGCGGA : 551
F9-Stop-5     : GGCCGAGGATGGTAATGGCCGAATGTGGGATGGCAAGCTTGGATTTTGGCCCTTTGGATGATGTTGGACAGGACAGAGAGAAATCCGAATACAGAGCCGGCAGGAGGCCCTGTTTGGGAAAGATGAGAGTGTGATAGAGACACTTAAAGGCGGA : 551
```

		*	1080	*	1110	*	1140	*	1170	*	1200	*			
F9-Stop-2	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGACAGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	700									
Aacutmar1.12	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	1213									
Aacutmar1.7	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	1213									
F9-Stop-1	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	703									
Aacutmar1.2	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	1213									
Aacutmar1.9	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	1213									
Aacutmar1.1	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	1213									
Aacutmar1.3	:	TACATGCTGGACAAGGTCATTCCCTGCCATTATAGACAAGTTCCCATGACATACTTGAAAAG	----	AGGGGTCAAAAATCCAAACAAGAGGGATCAACTGCTCATTTGGCTGATGGAGATGAAGAGTGGAAATGAAGCAATGGAGATACTCAAACG	:	1213									
F9-Stop-3	:	TATATGCTTGACAAGTAATTTTACGCCATCTTGAGAAAGTTTCTAATGCCGTGTTCTGTGAATGAGGGTGTGAAGATCCAGCAAGAGGTGCAAAATCCCATATTTCTGACAATGATGAAGAAATGGTTAGATGCCATGCAAAAGCATTTTACGA	:	694											
F9-Stop-4	:	TACATGCTTGACAAGGTCATTCCGCCCATCTTGAAAAGTTTCTGAATGCCATTCTGTGAATGAGGGTGTGAAGATCCAGCAGGATGGCGCCAAATCCCATATTTCTGACAATGATGAAGAAATGGTTAGAGGCCATGCAGAACATTTTATGA	:	703											
F9-Stop-5	:	TACATGCTTGACAAGGTCATTCCGCCCATCTTGAAAAGTTTCTGAATGCCATTCTGCGAAATGGAAGTGTGAAGATCCAGCAGGATGGCGCCAAATCCCATATTTCTGACAATGATGAAGAAATGGTTAGAGGCCATGCAGAACATTTTATGA	:	703											
			1230	*	1260	*	1290	*	1320	*	1350	*			
F9-Stop-2	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTCTCTATTATGAAGCTGCACCGACAACCTGACTTTGGCTCTCATCAAGGCAGTGGAGGATG	:	842										
Aacutmar1.12	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCCTCTATTATGAAGCTGCACCAACAACCGACTTTGGCTCTCATCAAGGCAGTGGAGGATG	:	1365										
Aacutmar1.7	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACCGAATTTGGCTCTCATCAAGGCAGTGGAGGATG	:	1365										
F9-Stop-1	:	TGAGAGCAGGATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACCGAATTTGGCTCTCATCAAGGCAGTGGAGGATG	:	855										
Aacutmar1.2	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACCGAATTTGGCTCTCATCAAGGCAGTGGAGGATG	:	1365										
Aacutmar1.9	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACCGAATTTGGCTCTCATCAAGGCAGTGGAGGATG	:	1365										
Aacutmar1.1	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACCGAATTTGGCTCTCATCAAGGCAGTGGAGGATG	:	1365										
Aacutmar1.3	:	TGAGAGCAGAATTAAACATCAAGCTGTACACCCAACCGGCTCAATCACC	GGACACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACCGAATTTGGCTCTCATCAAGGCAGTGGAGGATG	:	1365										
F9-Stop-3	:	TGAGTACAAGCTGAATATTCAGATATATAGTCAACCGGCTCAATCACC	GGATACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCAACAACGACTTTGGCTCTCATCAAGGCAGTGGAGGATG	:	846										
F9-Stop-4	:	CGAGTACAAGTTGAAGATTCAGATATACACTCAACCGGCTCAATCACC	GGATACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACGACTTTGGCTCTCATCAAGGCAGTGGAGGATG	:	855										
F9-Stop-5	:	CGAGTACAAGTTGAAGATTCAGATATACACTCAACCGGCTCAATCACC	GGATACAAACATCAATGACTTGGCATTCTTTCAGGAGCATCCAAAGCTCTTATTATGAAGCTGCACCGACAACGACTTTGGCTCTCATCAAGGCAGTGGAGGATG	:	855										
			1380	*	1410	*	1440	*	1470	*	1500	*			
F9-Stop-2	:	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	886
Aacutmar1.12	:	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	1517
Aacutmar1.7	:	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	1517
F9-Stop-1	:	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	899
Aacutmar1.2	:	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	1517
Aacutmar1.9	:	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	1517
Aacutmar1.1	:	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	1517
Aacutmar1.3	:	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	CATACTGGAAGTATCCAGCAAAACAAGCTCAATCAGATGTATTGA	-----	:	1517
F9-Stop-3	:	CATATTGGAAGTATCCAGTAAACAAGCTCAATCAGATGTATTGA	-----	CATATTGGAAGTATCCAGTAAACAAGCTCAATCAGATGTATTGA	-----	CATATTGGAAGTATCCAGTAAACAAGCTCAATCAGATGTATTGA	-----	CATATTGGAAGTATCCAGTAAACAAGCTCAATCAGATGTATTGA	-----	CATATTGGAAGTATCCAGTAAACAAGCTCAATCAGATGTATTGA	-----	CATATTGGAAGTATCCAGTAAACAAGCTCAATCAGATGTATTGA	-----	:	890
F9-Stop-4	:	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	:	898
F9-Stop-5	:	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	CGTATTGGAAGTATCCAAATTAACAAGCTCAATCAGATGTATTGA	-----	:	899

**Figure 76 : Alignement en acides nucléiques des 5 séquences ADNc de MLE (*F9-Stop-1* à *-5*) et des 6 séquences pleine taille et à cadre de lecture ouvert (*Aacutmar1*) isolées dans le présent travail et présentées dans le chapitre précédent.**



### **1.3. Détection par Q-RT-PCR de l'expression des MLE et des gènes de réponse aux stress**

Les résultats obtenus par la technique de RT-PCR n'ont pas été très concluant, aussi bien pour les gènes de ménage que pour les gènes de réponse aux stress et pour les MLE. Toutefois, le résultat positif obtenu pour les MLE après 5 h de stress à 16 °C et le fait que des éléments complets avec des ORF aient été mis en évidence, nous ont permis de définir d'autres amorces et d'utiliser la méthode de RT-PCR quantitative pour quantifier l'expression des MLE détectés.

Pour la PCR en temps réel, les ARN ont été extraits à l'aide d'un kit (Sigma) pour tous les échantillons et leur qualité testée par le système automatisé Experion. Les profils de microélectrophorèse ainsi qu'une valeur de RQI satisfaisante (8,2 en moyenne), synonyme d'une bonne qualité des ARN totaux ont été obtenus pour tous les échantillons (stage M2, Stanilas Thiriet-Rupert).

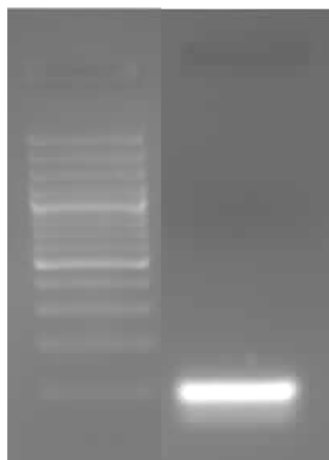
#### **1.3.1. Caractérisation des gènes de ménage**

Du fait du peu d'informations disponibles dans les banques de données sur *A. acutiuscula* NCC 216, ce travail s'est basé sur 12 gènes de ménage caractérisés chez la diatomée modèle *P. tricornutum* CCMP 632 (Siaut et al. 2007) ainsi que sur les deux gènes de ménage (actine et ribosome) caractérisés en RT-PCR chez *A. acutiuscula* NCC 216. Après plusieurs essais sur des ADNc obtenus à partir de cultures témoins (cultivées pendant 5 j à 24 °C), des amplifications ont été obtenues en utilisant les amorces ciblant 7 des gènes de ménages caractérisés chez *P. tricornutum* - actine 16, actine 12, GAPDH (Glyceraldéhyde 3-phosphate déshydrogénase), TBP (TATA box binding protéine), ubiquitine, RPS (protéine ribosomale de la petite sous unité 30S), H4 (Histone 4) - ainsi que pour l'actine  $\beta$  et la sous-unité ribosomale 18S d'*A. acutiuscula*. Le séquençage a confirmé l'amplification des gènes souhaités, excepté pour TBP et H4. Les 7 autres gènes de ménage ont donc été conservés pour la suite de ce travail (PCR en temps réel).

#### **1.3.2. Caractérisation des gènes d'intérêt**

Les ADNc utilisés pour caractériser ces gènes ont été obtenus à partir d'une culture témoin (cultivée 5 j à 24 °C). D'après le travail de Master 2 de S. Thiriet-Rupert, seul un gène de réponse aux stress a pu être conservé pour le reste de l'expérimentation, il s'agit du gène de hsp70 qui a pu être amplifié, cloné, séquencé et identifié comme tel.

Le couple d'amorces (*QHSPAmpfw*/*QHSPAmprw*) utilisé précédemment a été conservé pour la Q-RT-PCR (Fig. 77).



**Figure 77 : Électrophorèse sur gel d'agarose 1 % des produits d'amplification obtenus avec les amorces *QHSPAmpfw*/*QHSPAmprw* chez *A. acutiuscula*.**

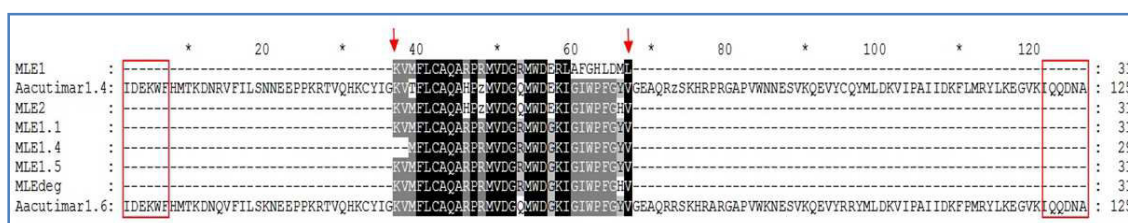
En ce qui concerne les MLE, dans le chapitre 1, deux nouvelles séquences de 400 pb ont été identifiées et ajoutées aux 6 précédentes déjà acquises au laboratoire (D. Hermann, soumises). Trois couples d'amorces (*acutimarF1*/*acutimarR1*; *acutimarF2*/*acutimarR2* ; *acutimarFdeg*/*acutimarRdeg*) dont un dégénéré ont été dessinés à partir de ces 8 séquences. Ces trois couples d'amorces ont permis l'amplification de 9 séquences dont 3 avaient une taille inférieure à celle attendue (48 pb et 66 pb au lieu de 95 pb). Les 6 séquences obtenues ont été notées MLE1, 1.1, 1.4, 1.5 pour celles issues de l'utilisation du couple F1/R1; MLE2 pour le couple F2/R2 et MLEdeg pour celle issue de l'amplification avec le couple d'amorces dégénérées. Les trois couples d'amorces ont été testés en PCR en temps réel.

Parmi les séquences de transposase de MLE obtenues précédemment, les deux plus proches (nommées *Aacutimar 1.4* et *1.6*) ont été utilisées dans le but de les comparer aux six séquences obtenues avec les amorces de PCR en temps réel. Au niveau nucléotidique, ces séquences ont un pourcentage d'identité entre elles variant de 87 à 99 % et allant de 87% à 99 % avec les séquences 1.4 (FN546188) et 1.6 (FN546190). Lorsque l'alignement a été refait sans les séquences des amorces, les séquences avaient un pourcentage d'identité entre elles compris entre 82% et 100 % (pour trois d'entre elles) et de 90 % à 100 % avec les séquences 1.4 et 1.6 (Fig. 78).



**Figure 78 : Alignement nucléotidique des six séquences obtenues avec les trois couples d'amorces ayant servi à la Q-RT-PCR et ciblant le gène de la transposase des MLE avec les deux séquences des banques de données les plus proches *Aacutimar1.4* et *1.6*.**

Toutes les séquences ont été traduites en protéines (grâce au logiciel Expasy Translate <http://web.expasy.org/translate/>) et leur alignement révèle l'apparition d'un codon STOP dans la séquence 1.4 ainsi que dans les séquences MLE1 et MLE2. En ce qui concerne les séquences MLE obtenues avec les amorces de Q-PCR, elles montrent, entre elles, un pourcentage d'identité de 83 % à 100 % et des pourcentages un peu plus faibles avec les séquences 1.4 et 1.6 de 81 % à 97 %. (Fig. 79).



**Figure 79 : Alignement protéique des six séquences obtenues avec les trois couples d'amorces utilisées en Q-RT-PCR et ciblant le gène de la transposase des MLE avec les deux séquences déposées dans les banques de données et obtenues dans thèse D. Hermann (2011).**

Les deux motifs caractéristiques comprenant les deux premiers acides aspartiques de la triade catalytique sont encadrés en rouge. Les codons stop (z) sont indiqués par une flèche rouge.

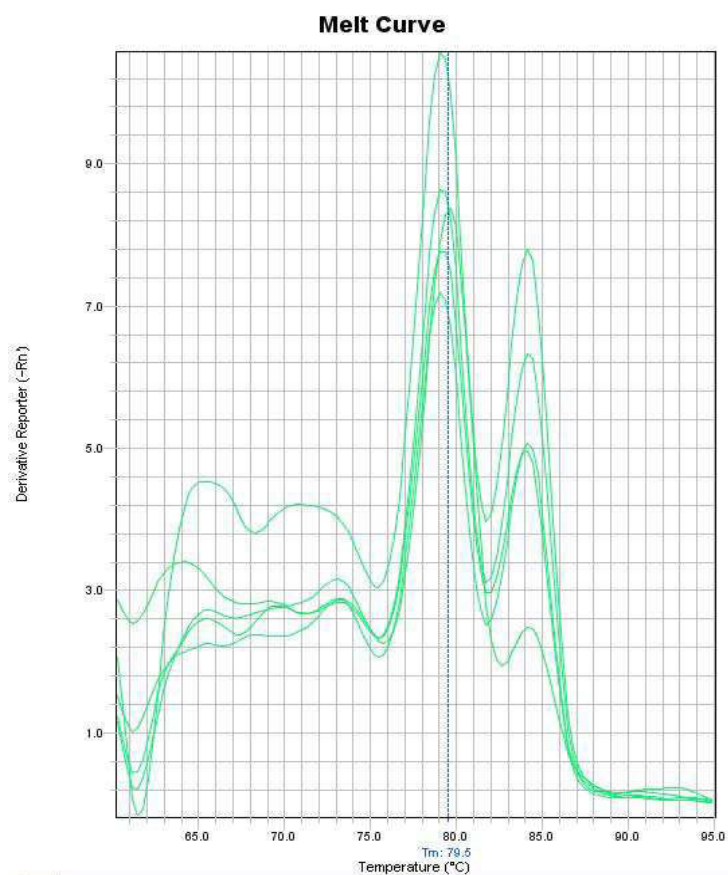
Il est à noter que la portion amplifiée par les amorces est comprise entre les deux premiers acides aspartiques de la triade catalytique au sein des motifs conservés (IDEKWF et IQQDNA) caractéristiques des MLE de plantes (Fig. 79).

### **1.3.3. PCR en temps réel**

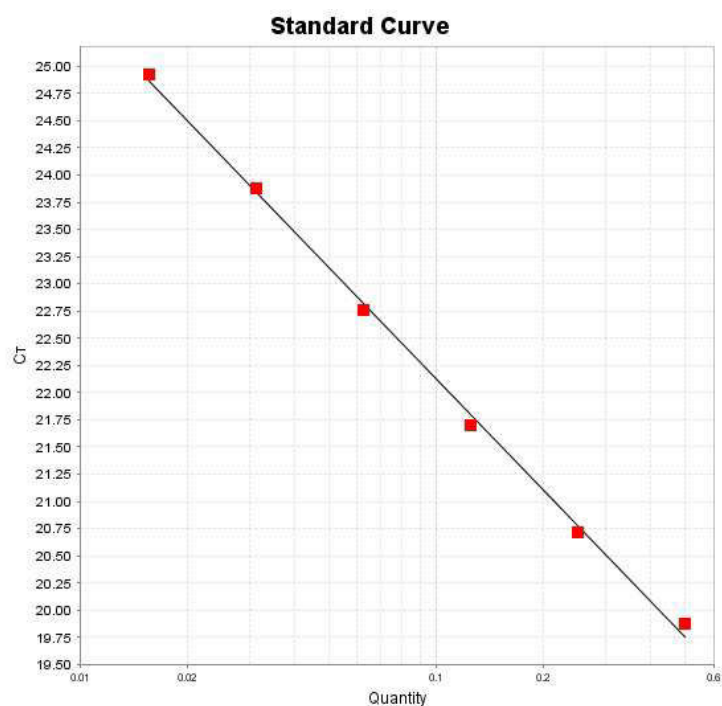
#### ***Mise au point***

Pour la mise au point en qPCR et ce, quelle que soient les amorces utilisées, la courbe de dissociation ne doit comporter qu'un seul pic. Si ce n'est pas le cas, cela signifie qu'il y a la présence de deux amplicons, et donc une amplification non spécifique générées par les amorces utilisées. Dans ce cas, elles doivent être changées. Ce fut le cas pour l'ubiquitine (Fig. 80).

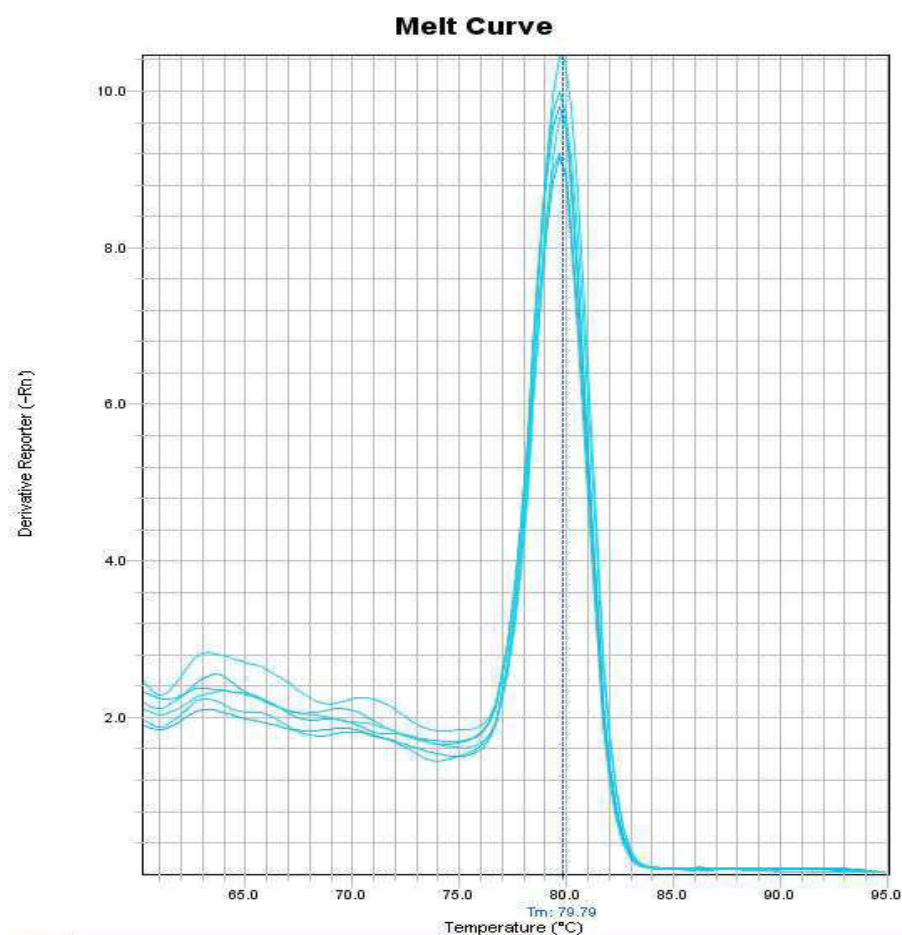
Ensuite, une courbe standard (logarithme de la concentration de matrice en fonction des valeurs de Ct) a été réalisée (Fig. 81) afin de déterminer l'efficacité des amorces utilisées. Pour cela, une amplification a été réalisée avec une gamme de dilution d'ADNc (1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128). L'efficacité a été calculée grâce à la pente de la droite obtenue et elle doit être comprise entre 85 % et 115 %. Le  $R^2$  doit, quant à lui, être supérieur à 0,98. Pour les amorces ciblant l'actine  $\beta$  et la sous unité ribosomale 18S ont été utilisées à une concentration finale de 0,1  $\mu$ M. Les efficacités obtenues étaient, dans un premier temps, de 117,6 % et 150,9 %. Les courbes de dissociation montrant une amplification d'un seul fragment (Fig. 82), le problème venait d'un biais de manipulation (le volume déposé dans chaque puits n'était pas régulier). Pour y remédier, une pipette multicanaux automatique a été utilisée et a permis, dans les mêmes conditions, d'obtenir des efficacités de 111,8 % pour l'actine  $\beta$  et de 107,5 % pour la sous unité ribosomale 18S (Tab. 18). Dans le cas de RPS, les amorces n'ont pas permis d'obtenir de courbe standard satisfaisante. Un nouveau couple d'amorces a donc été redéfini et, utilisé à une concentration finale de 0,8  $\mu$ M, a permis d'obtenir un  $R^2$  de 0,992 et une efficacité de 133,4 %. Afin d'améliorer cette dernière, les amorces ont été utilisées à une concentration finale de 0,5  $\mu$ M, ce qui a permis d'obtenir une efficacité de 105,6 % (Tab. 18). Quant aux amorces ciblant la GAPDH et l'actine 16, elles n'ont pas permis d'amplification. Suffisamment de gènes de ménage ayant déjà été mis au point (Tab. 18), aucun nouveau couple n'a été défini et ces deux gènes n'ont pas été conservés.



**Figure 80 : Courbe de dissociation obtenue avec le premier couple d’amorces ciblant le gène de l’ubiquitine. La présence de deux pics traduit celle de deux amplicons**



**Figure 81 : Courbe standard obtenue avec le deuxième couple d’amorces ciblant l’ubiquitine.**



**Figure 82 : Courbe de dissociation obtenue avec le couple d'amorces ciblant le gène 18S**

**Tableau 18 : Liste des couples d'amorces utilisés en PCR en temps réel et conditions retenues.**

Gène	Amorces	séquence	Efficacité	R <sup>2</sup>	[amorces] (μM)
Ubiquitine	<i>AacutUbifw</i> <i>AacutUbirv</i>	CGACAACGTCAAGACCAAGATTC CCATCTTCCAGCTGCTTTCC	96,999	0,998	0,8
RPS	<i>AacutRPSfw</i> <i>AacutRPSrv</i>	CCGTCGTGGAACAACAAATAGTC CTTCGACAAAGGCACCATAGG	105,677	1	0,5
Actine 12	<i>AacutActin12fw</i> <i>AacutActin12rv</i>	CGGAAGCGTACAAGGAAAGG CGAGAGCGTATGACGCAAATCATG	93,779	0,984	0,5
Actine β	<i>QPCRAacutActinbF</i> <i>QPCRAacutActinbR</i>	CTTGCCGGTCGCGATCT CGTTCAGCCGTAGTGGTCAA	111,841	0,996	0,1
18S	<i>QPCRAacut18SF</i> <i>QPCRAacut18SR</i>	AAAATTGGACCACAGGGAACAC CCAGCAGCCGCGGTAAT	107,538	0,997	0,1
HSP70	<i>QHSPAmpfw</i> <i>QHSPAmprv</i>	CATTAACGAACCAACTGCCGC CACATCAAATGTACCACCACCAAG	100,315	0,991	0,8
MLE	<i>acutimarF1</i> <i>acutimarR1</i>	AAAGTGATGTTCTTGTGTGC CCAACATATCCAAATGGCCAAAT	112,34	0,984	0,1

A une concentration de 0,1  $\mu\text{M}$ , le couple d'amorces retenu pour la hsp70 a montré une bonne spécificité d'amplification, un bon  $R^2$  mais une efficacité insuffisante (57,2 %). Afin de l'améliorer, la concentration finale des amorces a donc été augmentée à 0,5  $\mu\text{M}$  et 0,8  $\mu\text{M}$ . Ce qui a permis d'obtenir, respectivement, une efficacité de 94,5 % et 100,3 % (Tab. 18). Cette dernière condition a donc été retenue.

Pour les MLE, les trois couples ont été testés et ont tous montré une efficacité et un  $R^2$  satisfaisants respectivement compris entre 90 et 115 % et proche de 1 (Tab. 18). Cependant, la présence de deux pics sur la courbe de dissociation des couples acutimar F2/R2 et acutimar Fdeg/Rdeg, a révélé une amplification aspécifique. Le couple acutimar F1/R1 (112,34 % d'efficacité et  $R^2$  de 0,984) a donc été retenu pour la suite de ce travail.

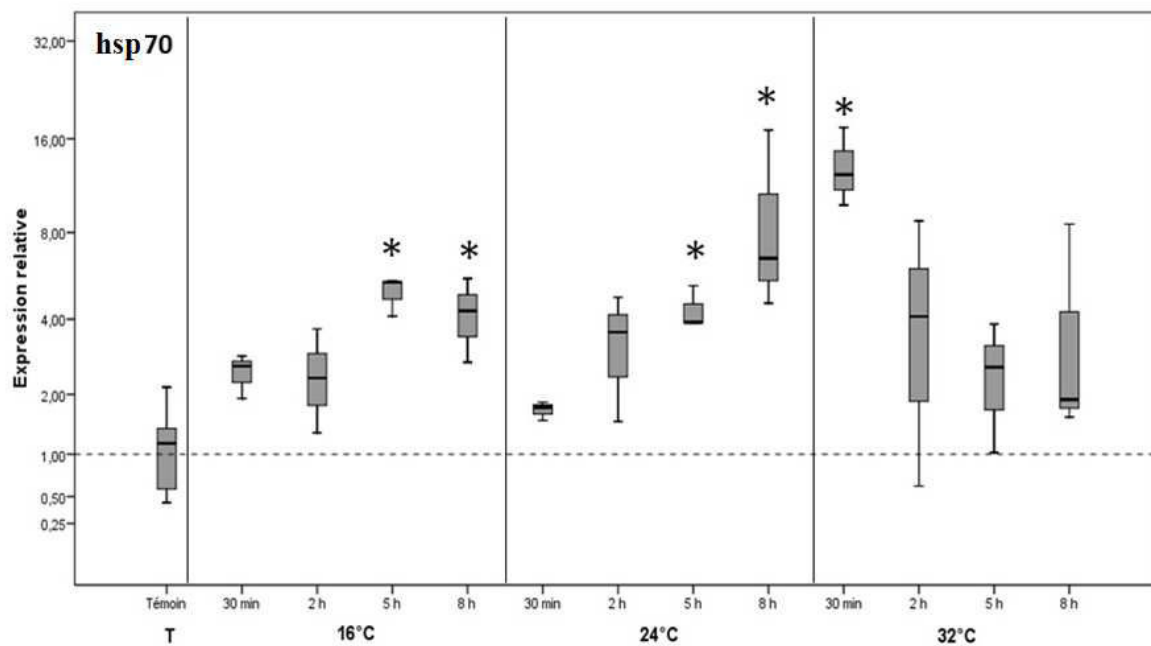
### ***Expression différentielle dans les différentes conditions thermiques***

Cette partie repose sur l'utilisation de la technique du  $2^{-\Delta\Delta C_t}$ . Chaque gène a été amplifié avec des ADNc obtenus à partir de cultures placées aux trois températures retenues, et dilués au  $\frac{1}{2}$ , excepté le gène de la sous unité ribosomale 18S dont les transcrits sont toujours plus abondants. Dans ce cas, les ADNc ont été dilués au  $\frac{1}{128}$ . Pour chaque condition, un triplicat biologique a été réalisé.

Les témoins ( $T=0$  cultivés à 24 °C) des conditions 16 °C et 24 °C ont été réunis afin de servir de référence aux valeurs de  $C_t$  des différentes conditions testées. Le témoin à 32 °C n'a pas été pris en compte car sa valeur semblait incohérente et l'échantillon a été éliminé. En effet, on observait une expression de presque 15 fois supérieure à celle du témoin pour cet échantillon alors que les diatomées n'avaient encore subi aucun changement de température. Ceci peut s'expliquer par le fait que le jour où ces 3 filtrations ont été réalisées, un défaut de climatisation de la salle de culture aurait engendré une température plus élevée que les conditions fixées pour l'expérience.

Les résultats obtenus (Fig. 83) montrent, dans le cas d'une exposition à une température de 16 °C durant 30 min et 2 h, que le niveau d'expression du gène de la hsp70 ne varie pas par rapport au témoin. En revanche, une surexpression est observée après 5 h (5 fois plus de transcrits) et 8 h (4 fois plus de transcrits). Le même effet est observé sur l'expression de ce gène suite à une exposition à une température de 24 °C pendant 5 h (4 fois plus exprimé) et 8 h (9 fois plus exprimé en moyenne). Dans ces conditions de température, nous notons que la quantité de transcrits tend à augmenter progressivement

au cours de l'expérimentation. En revanche, lors d'une exposition à une température de 32 °C, le gène de la hsp70 n'est surexprimé qu'après 30 min à cette température (13 fois plus exprimé). Il retrouve ensuite un niveau d'expression qui, bien que ne différant pas significativement du contrôle, montre une tendance à une diminution progressive au cours du temps.

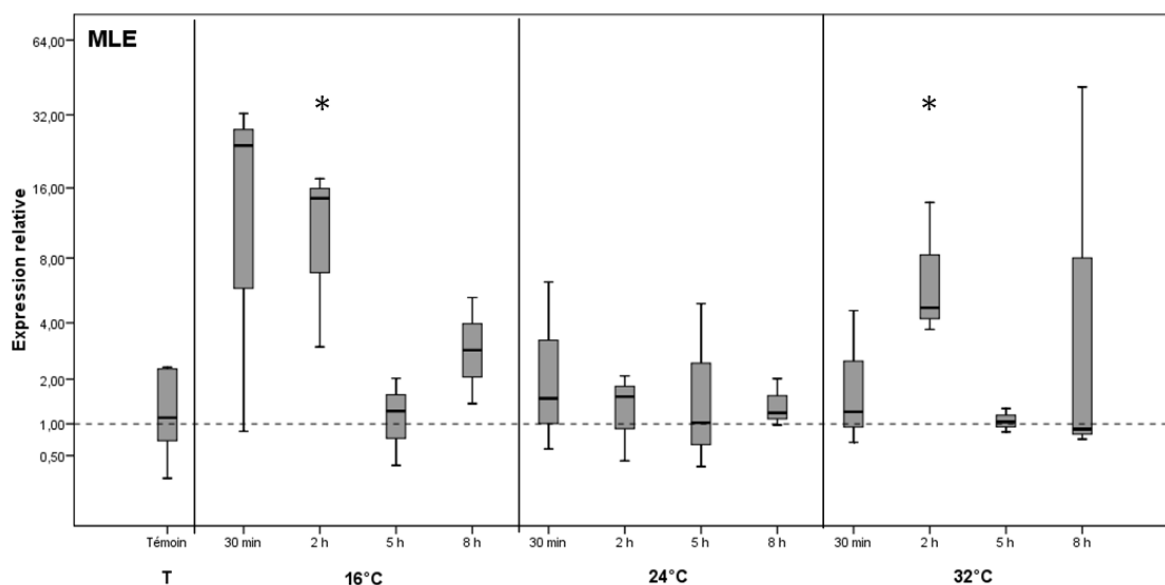


**Figure 83 : Expression relative du gène de la hsp70 chez *A. acutiuscula* placée pendant 0,5 à 8h à 16, 24 ou 32 °C comparativement au témoin (24 °C).**

Les étoiles (\*) montrent les résultats significatifs au seuil  $p < 0,05$  avec le test de Mann-Whitney sous SPSS par rapport au témoin



Pour ce qui est de l'expression du gène de la transposase des MLE (Fig. 84), il est observé une surexpression après une exposition de 2 h à une température de 16 °C (en moyenne 11 fois plus exprimé que dans le témoin) et potentiellement à 30 min (19 fois plus de transcrits en moyenne), mais ce dernier résultat n'est pas certain du fait de l'amplitude de la boîte à moustache ou box plot. Aucune variation d'expression n'est détectée dans le cas d'une exposition à 24 °C, qui est la température témoin, alors qu'après 2 h à 32 °C, on a observé, en moyenne, 7 fois plus de transcrits que dans le témoin. L'amplitude de variation de l'expression de la transposase après 8 h de stress à 32 °C ne permet pas de conclure définitivement quant au nombre de transcrits. Le profil d'expression tout au long de la cinétique étudiée pour cette température rappelle une expression transitoire déjà constatée en RT-PCR (Thèses D. Hermann, 2011 et FN. Egue 2013).



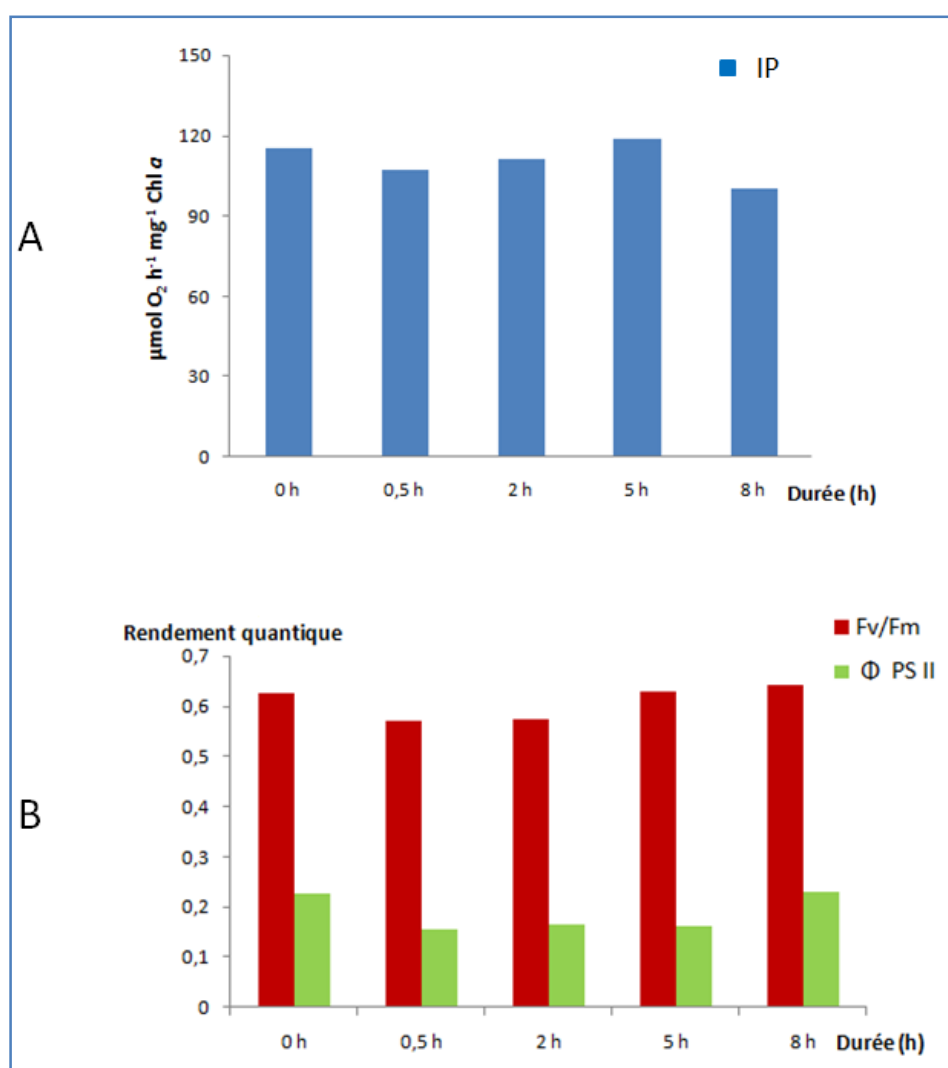
**Figure 84 : Expression relative du gène de la transposase des MLE chez *A. acutiuscula* placée pendant 0,5 à 8h à 16, 24 ou 32 °C comparativement au témoin (24 °C).**

Les étoiles (\*) montrent les résultats significatifs au seuil  $p < 0,05$  avec le test de Mann-Whitney sous SPSS par rapport au témoin.

## 2. Impact d'un excès de zinc

### 2.1. Activité photosynthétique

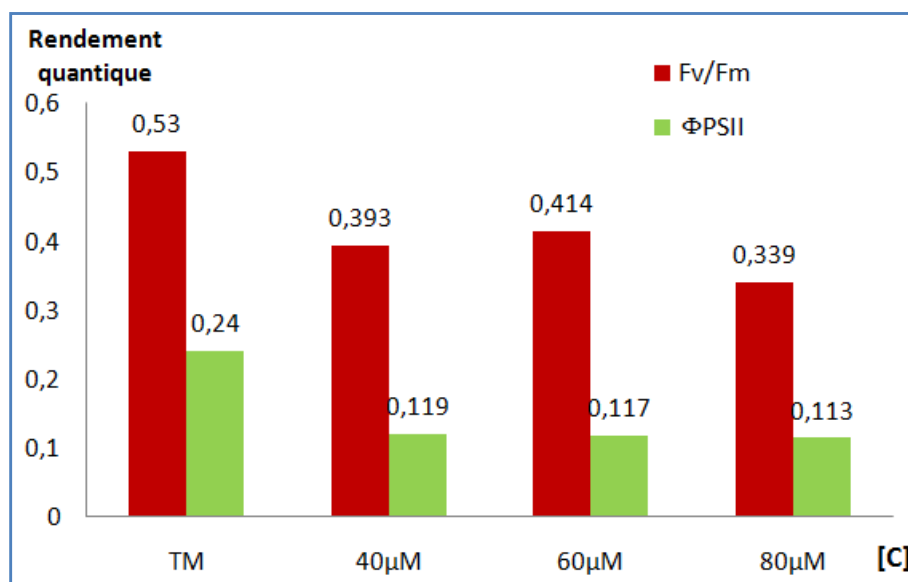
La diatomée *A. acutiuscula* a été cultivée sous un éclairage de 300  $\mu\text{mol}$  de photons  $\text{m}^{-2}\text{s}^{-1}$ , à 24 °C dans de l'eau de mer artificielle contenant une concentration en Zn de 0,25  $\mu\text{M}$ . A la fin de la phase exponentielle de croissance, du  $\text{ZnCl}_2$  20  $\mu\text{M}$  a été ajouté au milieu de culture. Puis après 0,5, 2, 5 et 8 h de contact avec cette forte concentration en Zn, l'activité photosynthétique a été mesurée et les ARN ont été extraits.



**Figure 85 : A-Évolution de l'intensité photosynthétique. B-Évolution du rendement quantique maximal ( $F_v/F_m$ ) et du rendement quantique effectif ( $\Phi \text{ PS II}$ ) du photosystème II chez *A. acutiuscula* en présence d'un excès de zinc (20  $\mu\text{M}$ ) pendant 0 à 8 h. (Les valeurs indiquées ci-dessus sont la moyenne de deux essais différents).**

Dans cet unique essai, l'IP n'est pas modifiée par l'ajout de zinc, par contre au niveau du PSII, le rendement quantique effectif tend à diminuer, et ce dès 30 min d'exposition au Zn ajouté (- 32 %). Fv/Fm diminue faiblement (- 9 %) pendant les deux premières heures d'exposition à l'excès de Zn puis retrouve le niveau des algues témoins. Cette concentration en zinc semble donc insuffisante pour perturber notablement le métabolisme de cette diatomée sur le court terme.

Craignant qu'il n'y ait pas de réponse visible par la suite au niveau moléculaire, nous avons testé 3 autres concentrations à savoir 40, 60 et 80  $\mu\text{M}$  en nous appuyant sur des doses là encore employées dans la littérature (Vavilin et al., 1998 ; Danilov et Ekelund, 2001). Une seule mesure des rendements quantiques a été réalisée après 8 h de contact afin de choisir la concentration à employer pour ensuite réaliser les échantillonnages pour la biologie moléculaire.



**Figure 86 : Évolution des rendements quantiques maximaux (Fv/Fm) et effectifs ( $\Phi\text{PSII}$ ) du photosystème II chez *A. acutiuscula* placée pendant 8 h dans de l'eau de mer enrichie en zinc.**

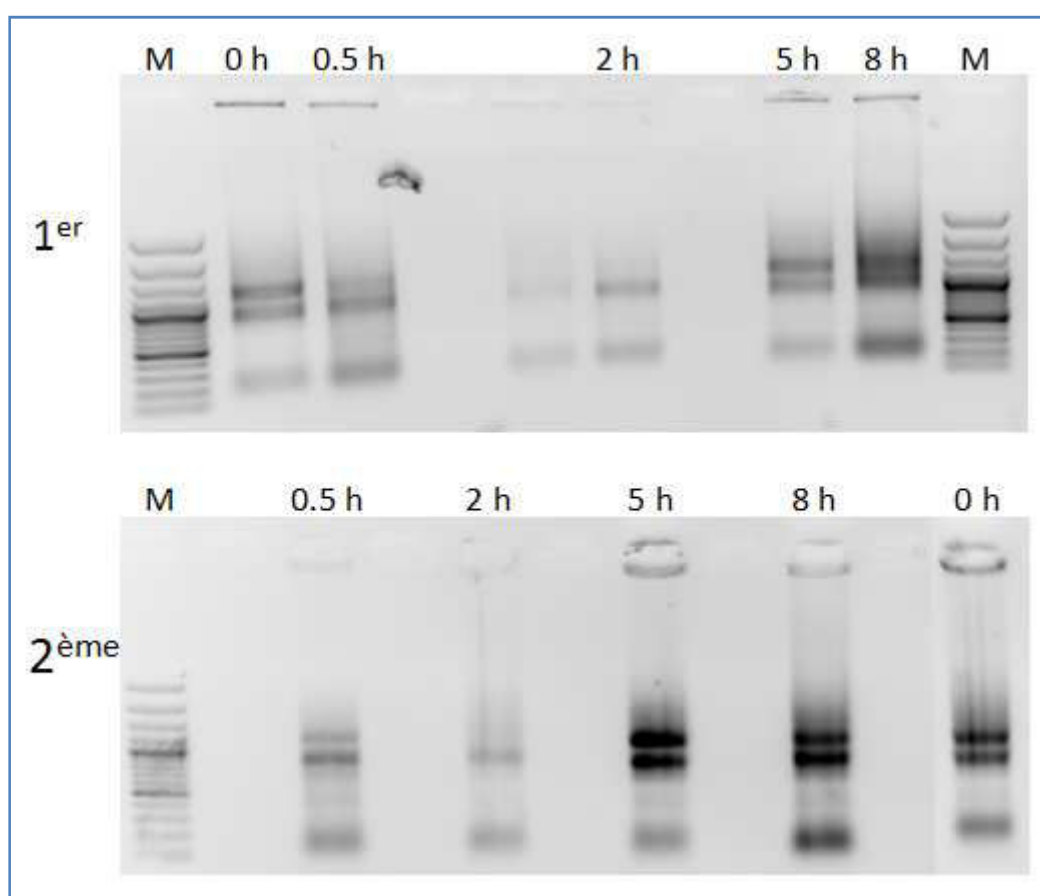
Concentration en zinc ajouté : 40  $\mu\text{M}$ , 60  $\mu\text{M}$  et 80  $\mu\text{M}$ . TM : témoin (microalgues cultivées dans l'eau de mer non enrichie en zinc).

La figure 86 montre qu'avec 40  $\mu\text{M}$  de Zn, les rendements quantiques maximum (Fv/Fm) et effectifs ( $\Phi\text{PSII}$ ) sont nettement réduits par rapport à ceux des diatomées témoins. Nous avons donc choisi cette concentration pour tester l'impact d'un excès de zinc au niveau moléculaire puisque la photosynthèse est affectée (réduction de 50 % de  $\Phi\text{PSII}$ ) durablement (8 h) sans être complètement inhibée.

## **2.2. Détection de l'expression des MLE**

### **2.2.1. Extraction des ARN totaux**

Les extractions d'ARN totaux réalisées à partir des diatomées exposées pendant 0,5 à 8 h au zinc 40  $\mu$ M sont de moins bonne qualité que ceux issus des traitements thermiques. Toutefois, sur la figure 87, nous voyons les deux bandes ribosomales 28 et 18S avec parfois une légère dégradation du 28S, notamment après 2 h de stress. Tous les réplicats ont présenté une qualité moyenne des ARN qui n'a néanmoins pas gêné les expérimentations suivantes.



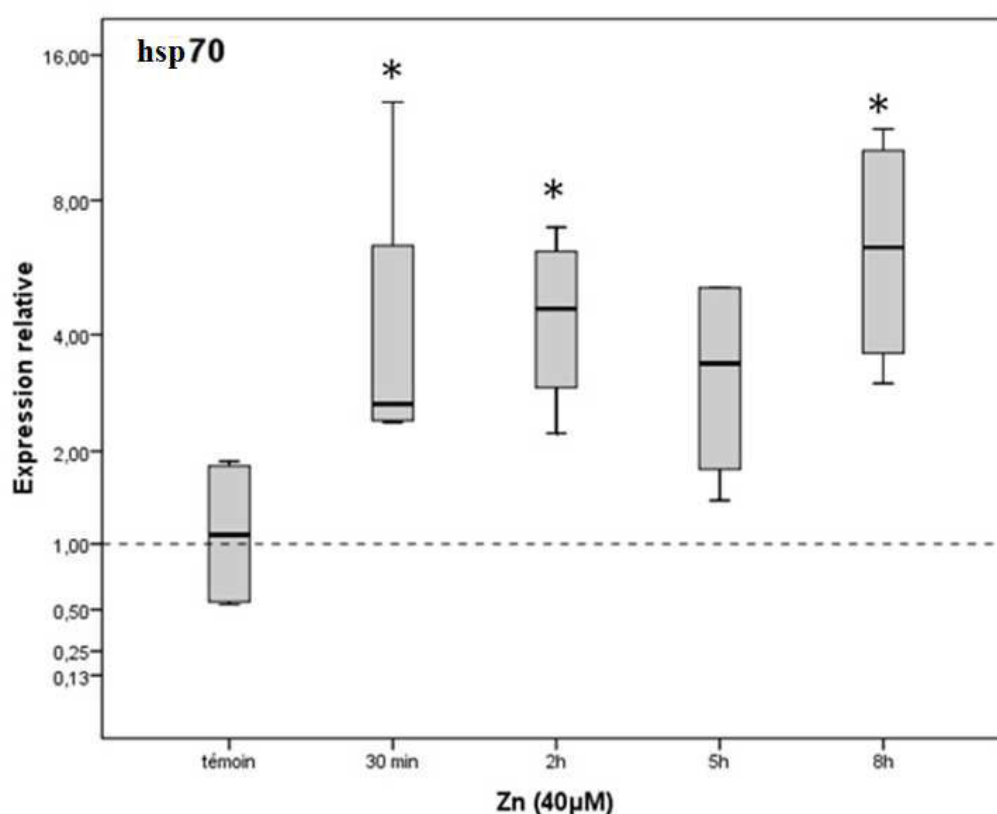
**Figure 87 : Électrophorèse des ARN totaux extraits d'*A. acutiuscula* exposée pendant 0,5 à 8 h à un excès de zinc (40  $\mu$ M). 1<sup>er</sup> et 2<sup>ème</sup> réplicats.**

### 2.2.2. Q-RT-PCR

Ayant obtenu, en stress thermique des résultats concluants en Q-RT-PCR et peu concluants en RT-PCR classique, nous avons décidé de n'exploiter que la technique quantitative pour les cultures exposées à un excès de zinc.

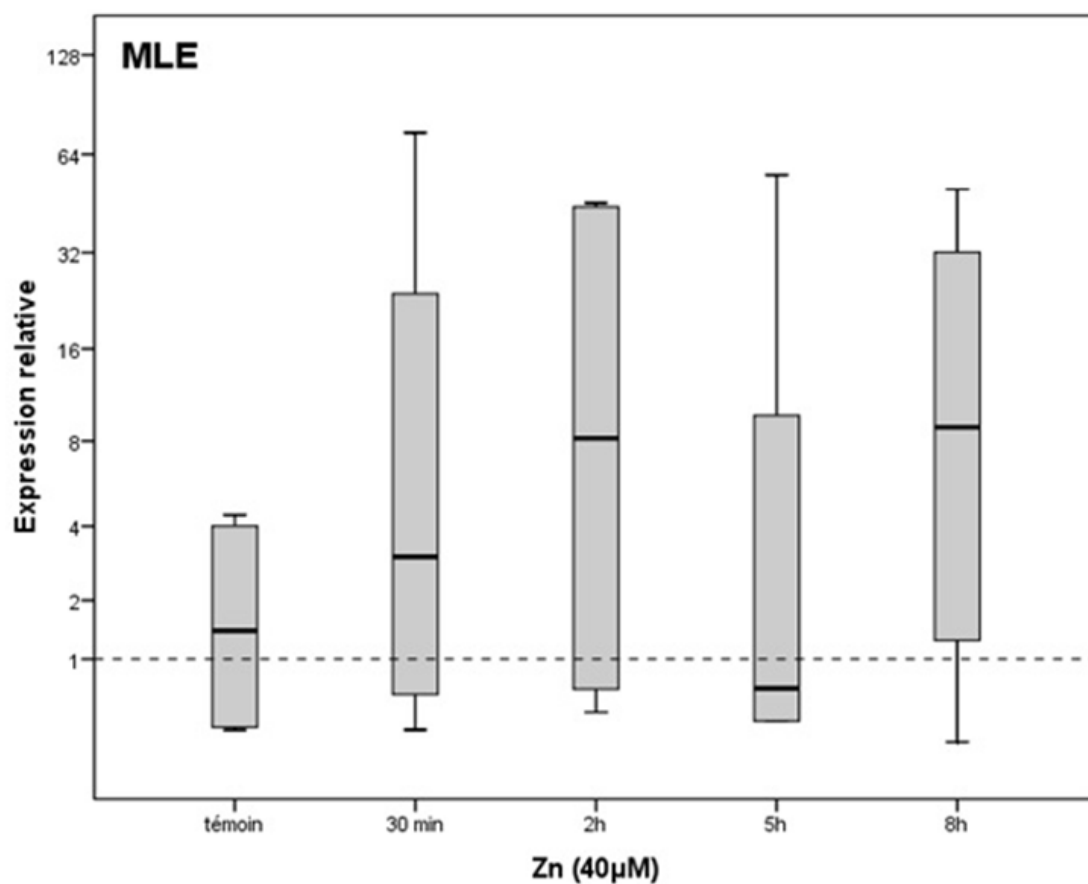
Les résultats obtenus montrent globalement une sur-expression comprise entre trois et 15 fois pour le gène de hsp70 (Fig. 88) à 30 min, 2 h et 8 h d'exposition à 40  $\mu$ M de zinc. Après 5 h, en revanche, aucune différence avec le témoin n'a été détectée mais une tendance à la sur-expression est possible du fait de la place de la médiane dans la boîte à moustache.

En ce qui concerne les MLE, une grande variation du niveau d'expression du gène de la transposase pour un même temps d'exposition a été observée (Fig. 89), notamment après 8 h de stress. Toutefois de manière générale, l'expression est restée stable au cours de la cinétique de stress avec une tendance à la sur-expression, si l'on regarde les médianes.



**Figure 88 : Expression relative du gène de la hsp70 chez *A. acutiuscula* exposée pendant 0,5 à 8 h à un excès de zinc (40  $\mu$ M) dans le milieu de culture comparativement au témoin.**

Les étoiles (\*) montrent les résultats significatifs au seuil  $p < 0,05$  avec le test de Mann-Whitney sous SPSS par rapport au témoin.



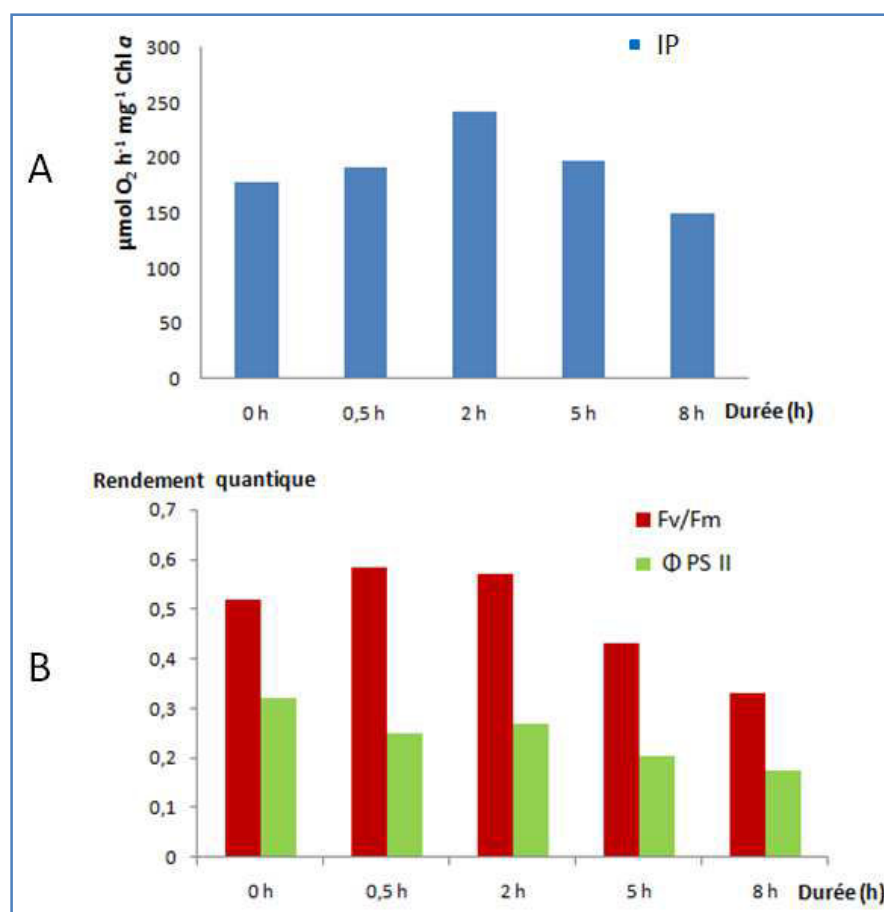
**Figure 89 : Expression relative du gène de la transposase des MLE chez *A. acutiuscula* exposée pendant 0,5 à 8 h à un excès de zinc (40 µM) dans le milieu de culture comparativement au témoin.**

Les étoiles (\*) montrent les résultats significatifs au seuil  $p < 0,05$  avec le test de Mann-Whitney sous SPSS par rapport au témoin.

### **3. Impact d'un excès de cuivre**

#### **3.1. Activité photosynthétique**

La diatomée *A. acutiuscula* a été cultivée sous un éclairage de  $300 \mu\text{mol photons m}^{-2}\text{s}^{-1}$  à  $24^\circ\text{C}$  dans de l'eau de mer artificielle contenant une concentration en Cu de  $0,15 \mu\text{M}$ . Afin de déterminer l'impact d'un excès de cuivre, en fin de phase exponentielle de croissance, du  $\text{CuCl}_2$   $3 \mu\text{M}$  a été ajouté au milieu de culture. L'activité photosynthétique a été mesurée après 0, 0,5, 2, 5 et 8 h de ce traitement (Fig. 90).

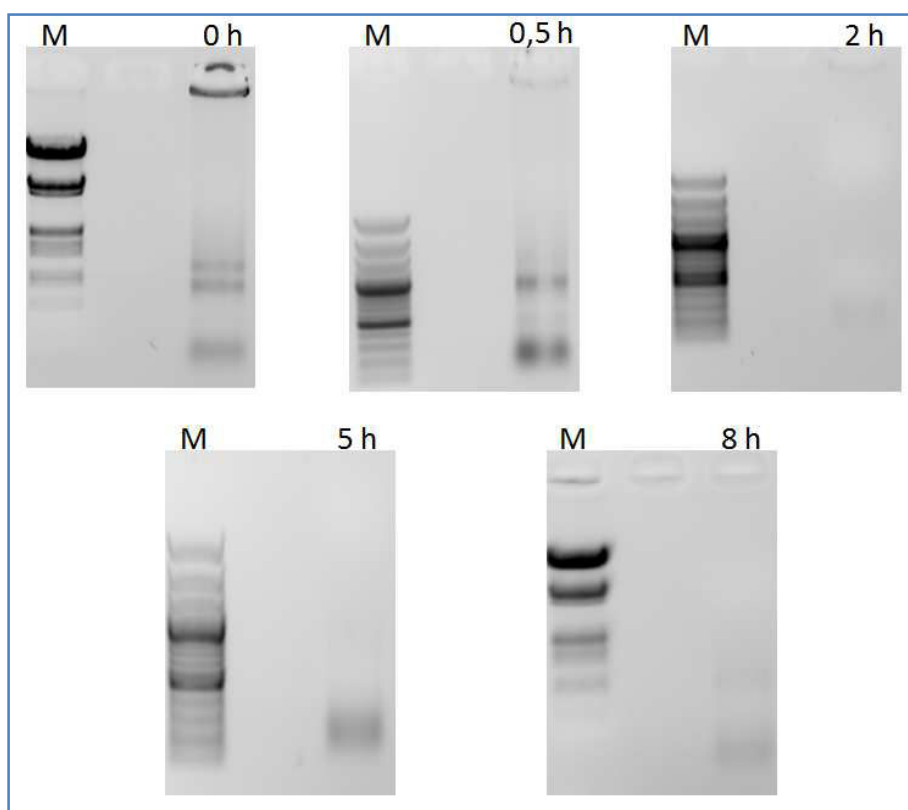


**Figure 90 : A : Évolution de l'intensité photosynthétique. B : Évolution du rendement quantiques maximal ( $F_v/F_m$ ) et du rendement quantique effectif ( $\Phi_{PSII}$ ) du photosystème II chez *A. acutiuscula* en présence de cuivre  $3 \mu\text{M}$  pendant 0 à 8 heures. Expérience réalisée une seule fois.**

L'ajout de cuivre a augmenté faiblement l'IP de 36 % et  $F_v/F_m$  de 10 % pendant les deux premières heures de traitement puis ces valeurs diminuent progressivement jusqu'à la fin de l'expérimentation. Une diminution régulière de  $\Phi_{PSII}$  est observée pendant toute la durée d'exposition à l'excès de Cu, cette diminution est de 22 % après 30 min de traitement et atteint 45 % après 8 h de traitement.

### **3.2. Détection de l'expression des MLE**

De la même façon que précédemment, les ARN totaux ont été extraits de culots provenant de cultures de microalgues cultivées en présence de 3  $\mu$ M de cuivre. Les résultats de ces extractions sont présentés en figure 91. Comme le montrent les gels d'agarose, les profils d'ARN sont extrêmement dégradés. De nombreux tests, ici non présentés, ont été réalisés pour essayer d'améliorer les résultats : changement de tous les produits utilisés, utilisation de matériel passé au four Pasteur, extraction directement après récolte sans stockage à -80 °C, utilisation du kit d'extraction, etc. Aucune amélioration n'a été obtenue. Toutefois des tests de RT-PCR ont tout de même été tentés sans résultats. Ces échantillons n'ont donc pas été testés en PCR quantitative. De nouveaux essais d'extraction seront à faire pour permettre l'exploitation de ces données. Il pourrait par exemple s'agir de rinçage les microalgues avec du PBS 1 % afin d'éliminer plus de Cu fixé aux polysaccharides qui entourent le frustule et qui pourrait perturber l'extraction.



**Figure 91 : Électrophorèse des ARN totaux extraits d'*A. acutiuscula* exposée à un excès de cuivre (3  $\mu$ M) pendant 0,5 à 8 h (exemple d'un réplicat).**

M : marquer de taille.



**Table 19 : Récapitulatif des réponses significativement différentes de celles des cultures témoins observées chez *A. acutiuscula* exposée à différentes températures ou à un excès de zinc pendant 0,5 à 8 h.(voir tableaux 15, 16, 17; figures 83, 84, 85, 86 88 et 89). Non m.: paramètre non mesuré.**

Traitement	Durée (h)	IP	Fv/Fm	$\Phi$ PSII	hsp70	MLE
16°C	0,5	X				
	2					X
	5	X			X	
	8	X			X	
24 °C	0,5					
	2					
	5				X	
	8				X	
32 °C	0,5			X	X	
	2					X
	5					
	8					
Zn 40 mM	0,5	Non m.	1 essai	1 essai	X	
	2	Non m.	1 essai	1 essai	X	
	5	Non m.	1 essai	1 essai		
	8	Non m.	1 essai	1 essai	X	

## **4. Discussion**

### **4.1. Effets des variations de température**

L'activité photosynthétique est l'une des réactions du métabolisme primaire des diatomées. Nos résultats montrent que l'IP est peu ou pas modifiée lorsque les microalgues subissent une augmentation de température de 8 °C (24 à 32 °C). Par contre, un abaissement de température de 8 °C (24 à 16 °C) a affecté l'IP de *A. acutiuscula*, cette réponse de l'appareil photosynthétique est significative dès 30 min à 16 °C. Par conséquent, cette diatomée tolère mieux une augmentation de température qu'un abaissement de température. Selon l'origine des diatomées, l'optimum thermique pour leur activité photosynthétique est différent. Le taux relatif de transport des électrons des diatomées des régions tempérées est maximal entre 25 °C et 35 °C (Yun et al., 2010), tandis que celui des diatomées d'origine tropicale est maximal à des températures plus élevées. Le traitement à 32 °C n'induit pas de modification importante de l'activité photosynthétique de la diatomée *A. acutiuscula* sauf pendant la première demi-heure pendant laquelle l'IP est augmentée sans être significativement différente de l'IP des algues témoins. Cette augmentation transitoire pourrait être due à une augmentation de la vitesse des réactions enzymatiques, notamment celle de la RuBisCO, et/ou du flux des électrons dans les chloroplastes. Chez les plantes, une augmentation modérée de la température stimule le flux d'électrons, l'activité oxygénasique de la RuBisCO et ainsi la production de H<sub>2</sub>O<sub>2</sub> (Sharkey, 2005).

A 16 °C, l'IP d'*A. acutiuscula* est nettement diminuée par rapport à celle mesurée à la température de culture habituelle de 24 °C. Ceci indique que cette diatomée est très sensible au froid, le traitement à 16 °C peut être considéré comme un stress thermique pour cette espèce. Le froid a pour effet de rigidifier les lipides membranaires et le cytosquelette des cellules, et de ralentir les activités enzymatiques. Nishida et Murata (1996) ont observé chez le tabac et des cyanobactéries que les membranes passent d'un état fluide liquide-cristallin à celui de type gel-solide lorsque la température diminue. Ce changement d'état modifie la perméabilité membranaire, les échanges ioniques diminuent et la photosynthèse est affectée.

Les paramètres de la fluorescence de la chlorophylle fournissent des informations précises sur l'état du PSII (Maxwell et Johnson, 2000). Le niveau de fluorescence dépend des deux autres processus de dissipation de l'énergie, la photochimie et l'émission de chaleur.

Fv/Fm représente le rendement quantique maximal de la photochimie du PSII. Les échantillons témoins d'*A. acutiuscula* ont une valeur moyenne de Fv/Fm de 0,6. Cette valeur indique qu'en moyenne 60 % des photons absorbés par les microalgues sont utilisés pour la photochimie. Ce rendement quantique moyen est un indicateur de l'état physiologique des microalgues, il est similaire à celui d'autres diatomées (*P. tricornutum* et *Odontella aurita*) non stressées (entre 0,6 - 0,7) (Nymark et al., 2009; Guihéneuf et al., 2010) et inférieur à celui des plantes (entre 0,7 - 0,8) (Dongsansuk et al., 2013). Les différents traitements thermiques que nous avons appliqués à *A. acutiuscula* n'ont pas modifié significativement ce paramètre ; un résultat comparable été obtenu chez une plante tropicale (*Musa* sp.) et une plante de région tempérée (*Allium ursinum*) soumises à une gamme de températures allant de -15 à 60 °C (Dongsansuk et al., 2013).

Le rendement quantique effectif  $\Phi$ PSII permet de déterminer l'efficacité du fonctionnement du PSII à la lumière. L'augmentation importante de  $\Phi$ PSII après 30 min d'exposition à 32 °C indique une optimisation transitoire de l'activité photosynthétique. Elle est peut être liée à une modification de l'état redox des intermédiaires (les quinones Q<sub>A</sub> et Q<sub>B</sub>, plastoquinone et plastocyanine) entre les deux PS et/ou une accélération du flux d'électrons due à une plus grande fluidité des membranes des thylacoïdes (Jupin et Lamant 1997 ; Oxborough, 2004 ; Schrader et al. 2004). La stabilisation de  $\Phi$ PSII à partir de 2 h d'exposition à 32 °C et ce, jusqu'à la fin du traitement, indique que de l'appareil photosynthétique s'est acclimaté. La capacité du PSII à retrouver son niveau de fonctionnement initial est une des bases de l'acclimatation (Allakhverdiev et al., 2008), les dommages au PSII chez les plantes ne se produisent qu'à des températures très élevées, généralement supérieures à 45°C (Sharkey, 2005).

Nous pouvons constater que les variations de l'IP et de  $\Phi$ PSII ne sont pas toujours étroitement corrélées. En effet, à 16 °C, l'IP décroît régulièrement et significativement au cours des 8 h de traitement alors que  $\Phi$ PSII diminue plus faiblement. De même à 32 °C, l'augmentation de IP est plus marquée que celle de  $\Phi$ PSII. La vitesse des processus mesurés, émission de O<sub>2</sub> (réaction lente par rapport à l'émission de fluorescence) et fluorescence de la Chl *a* (phénomène très rapide), peuvent expliquer ces différences. Plusieurs travaux (Shubert et al., 1995 ; Saakov, 2002 ; Rohacek et al., 2008 ; Lavaud et Lepetit, 2013) ont montré que les complexes protéines-pigments des antennes collectrices chez les microalgues et les plantes réagissaient très rapidement

aux variations de l'environnement et contribuaient ainsi à protéger l'appareil photosynthétique.

Sur le plus long terme, la synthèse de protéines et d'autres molécules peut être modulée selon les conditions environnementales et les espèces. La quantification des transcrits de quelques gènes de la diatomée marine *Thalassiosira pseudonana* cultivée à 12 ou 22 °C sous un éclairage de 50 ou 300  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  a montré que sous fort éclairage et à 12 °C, les transcrits du gène de la sédoheptulose 1,7-bisphosphate du cycle de Calvin sont moins abondants alors que ceux des gènes impliqués dans la photorespiration et la réduction des nitrates sont plus abondants qu'à 22 °C (Parker et Armbrust, 2005). Ces deux dernières voies métaboliques permettraient de dissiper l'énergie lumineuse en excès en diminuant le flux de carbone dans le cycle de Calvin, la sédoheptulose 1,7-bisphosphate étant inhibée à basse température. Cette espèce possède une capacité à régler finement les réponses d'acclimatation à des déséquilibres énergétiques entre la photochimie, la photorespiration et le métabolisme de l'azote (Parker et Armbrust, 2005).

Au cours de cette étude, l'effet d'une augmentation (32 °C) ou d'une diminution (16 °C) de température a été évalué au cours d'une cinétique de 8 h, grâce à l'expression du gène de hsp70 et de celui de la transposase de MLE. Les résultats obtenus montrent une réponse moléculaire différente en fonction du traitement, avec une légère sur-expression des hsp70 et des MLE lors du stress thermique à 16 °C alors que l'on observe une plus forte réponse des hsp70 lors de l'excès de zinc parallèlement à un niveau de transcrits stable des MLE.

De nombreuses hsp70 ont été identifiées chez les végétaux (Boston et al., 1996; Vierling, 1991). Les hsp70 sont désormais connues pour être exprimées lors de différents types de stress environnementaux comme les variations thermiques (hautes ou basses), la sécheresse, ou encore lors de stress chimiques (Lin et al., 2001; Sung et al., 2001; Guy et Li, 1998). Les diatomées, comme les autres organismes sont capables de répondre et de s'adapter à des stress environnementaux grâce, elles aussi, aux hsp (Rajadurai et al., 2005). Bonin et al. (1986) ont montré que les diatomées avaient su exploiter un grand nombre d'habitats grâce à leur capacité à s'adapter à différentes conditions environnementales et en particulier à une large gamme de températures. Toutefois, il existerait des exceptions, puisque Rajadurai et al. (2005) ont montré par des expérimentations d'acclimatation, que la diatomée *Amphora coffeaeformis*

était capable de croître sous une large gamme de températures sans pour autant s'adapter véritablement à des températures supérieures à 28 °C. *A. coffeaeformis* a été cultivée pendant plusieurs jours à 28 °C qui est la température témoin ou à 38 °C (température haute), puis placée à 42 °C. Dans les deux cas, la croissance est identique, la diatomée réagissant de la même façon avec ou sans acclimatation préalable à 38 °C. Plus récemment, les différences existant au niveau de l'expression des hsp d'une diatomée thermo-intolérante (*Phaeodactylum tricornutum*) et d'une thermo-tolérante (*Chaetoceros muelleri*) a été mis en évidence par Western-blot et SDS-PAGE (Rousch et al., 2004). Les auteurs ont montré que les cellules de diatomées thermo-tolérantes contenaient des niveaux relativement élevés de hsp.

Dans notre expérimentation, l'exposition à une température de 32 °C montre une forte expression du gène codant la hsp70 dès 30 minutes de stress (de 13 fois supérieure à celle du témoin). Le nombre de transcrits diminue ensuite progressivement au cours de la cinétique. Ce type de cinétique d'expression au cours du temps est en accord avec celles obtenues chez *Arabidopsis thaliana* (Sung et al. 2001) ou encore *Pisum sativum* (Lee et Vierling, 2000). Une sur-expression précoce des hsp70 permettrait aux microalgues de s'acclimater au choc thermique auquel elles sont exposées en prévenant d'éventuels dommages et agrégations de leurs protéines ainsi qu'en préservant la fluidité membranaire (Tsvetkova et al., 2002) et le photosystème II (Neta-sharir et al, 2005). La diminution progressive, mais néanmoins rapide, du nombre de transcrits pourrait s'expliquer par le fait que la souche étudiée est originaire d'une région du sud-Vietnam, où des températures élevées sont assez fréquentes. De plus, ces diatomées se retrouvent souvent dans des flaques lorsque la mer se retire. Ce type de biotope est sujet à de fortes augmentations de température puisque la surface et la profondeur de l'eau sont restreintes et ne permettent pas une diffusion de la chaleur et de l'énergie reçue. Ces microalgues auraient donc une plasticité génétique leur permettant de s'acclimater plus facilement à des températures plus élevées. Cette plasticité peut expliquer l'amplitude des valeurs obtenues après 2 h, 5 h et 8 h d'exposition à 32 °C. En effet, toutes les microalgues subissant le stress ne réagissent pas forcément en même temps ni avec la même intensité. Chez certaines, le nombre de transcrits peut diminuer plus rapidement que chez d'autres. Les résultats obtenus peuvent donc varier entre les différents réplicats. Nous pouvons également penser que, lors de cette diminution du nombre de transcrits de hsp70, une autre hsp soit produite pour faire face à ces hautes températures.

Des études chez les plantes *Arabidopsis thaliana* (Forreiter et al., 1997) et *Pisum sativum* (Lee et Vierling, 2000) ont mis en évidence une collaboration de hsp70 avec les small hsp (shsp) dans la prévention de la dénaturation des protéines. De plus, il a été démontré que la chaperonne hsp70 joue un rôle dans la régulation de la réponse au stress thermique et notamment lors de l'atténuation de celle-ci (comme le montre le profil d'expression obtenu) en interagissant avec le domaine de transactivation de facteurs de transcription impliqués dans la réponse au stress thermique, les Heat Shock Transcription Factors (HSTF) (Shi et al., 1998). Ceci entraîne la répression de la transcription des gènes de réponse au stress thermique. Les hsp70 peuvent également être impliquées dans la fonctionnalisation (multimérisation) de HSTF comme chez la microalgue *Chlamydomonas reinhardtii* (Schulz-Raffelt et al., 2007).

La mesure de la réponse moléculaire est à mettre en relation avec la réponse physiologique des individus. Une exposition à 32 °C pendant 30 min augmente significativement le rendement quantique effectif du PSII puis la diatomée s'acclimate puisque ce paramètre retrouve ensuite un niveau proche du témoin. Cette hypothèse est renforcée par le fait que l'intensité photosynthétique ne subit pas de variation significative lors de ce traitement.

L'exposition à la température témoin de 24 °C entraîne une expression inattendue du gène de la hsp70 après 5 h et 8 h de stress. Celle-ci montre un stress qui peut être imputé au déplacement de l'erlenmeyer jusqu'au cryothermostat ou à une différence de température entre l'air de la salle culture et celle de l'eau du cryothermostat, toutes deux étant pourtant en principe à 24 °C. Ces protéines étant impliquées dans de nombreux mécanismes cellulaires, elles sont exprimées de manières différentes tout au long de la vie des organismes. De plus, les hsp70 jouent un rôle dans la réponse au stress oxydatif chez les végétaux (Scarpeci et al., 2008) et il a été démontré que certains HSTF jouaient un rôle dans la détection des espèces réactives de l'oxygène ainsi que dans la régulation de l'expression de gènes de réponse au stress oxydatif chez *A. thaliana* (Miller et Mittler, 2006) et notamment l'ascorbate peroxydase (Panchuk et al., 2002). Or, dans cette étude, aucun lien ne peut être fait entre le stress oxydatif et la production de hsp70 puisque la MnSOD n'a pu être utilisée. Effectivement, bien que la MnSOD soit présente dans le génome d'*Amphora acutiuscula*, l'expression de ce gène n'a pas pu être détectée. L'isoforme manganèse n'est peut être pas sollicitée lors d'un stress chez cette microalgue comme chez la macroalgue *Ulva fasciata* lors d'un stress au cadmium où les

transcrits de l'isoforme à fer qui augmentent au détriment de ceux à manganèse (Wu et al., 2009).

Dans le cas d'une exposition à une température de 16 °C, une sur-expression du gène de *hsp70* est observée au bout de 5 h et 8 h de stress, plus tardivement que lors d'une élévation de température. Un tel délai dans la mise en place de la réponse à ce traitement peut s'expliquer par le fait que les basses températures induisent des changements moins brusques et moins intenses que de hautes températures. Pourtant, lors d'une exposition à une température de 16 °C, l'intensité photosynthétique subit une forte baisse dès 30 min et tout au long de la cinétique de stress. Ceci démontre un stress important même s'il n'entraîne pas immédiatement l'expression du gène de *hsp70*.

En RT-PCR, l'amplification d'un fragment de MLE n'a pas été aisée. Un fragment de MLE a pu être mis en évidence après 5 h d'exposition à 16 °C et pas dans les autres conditions. La technique reste peu sensible pour des gènes peu exprimés. Toutefois, ce résultat a été encourageant et confirmait les résultats obtenus lors d'une thèse précédente où une expression avait été mise en évidence à ce même temps sur un réplicat (Thèse D. Hermann, 2011). En RT-PCR-quantitative, l'expression du gène codant la transposase des MLE-éléments dont la transposition semble être influencée par la température (Giraud et Capy, 1996) - il n'est observé aucune modification du nombre de transcrits au cours du temps lors d'une exposition à la température témoin de 24 °C, alors qu'après 2 h à 32 °C, 13 fois plus de transcrits sont détectés. L'amplitude de variation de l'expression après 8 h d'exposition à 32 °C étant due à une seule valeur extrême, il serait bon de refaire des tests pour ce point. Ayant très peu de recul face à une telle étude, tant du point de vue du gène cible (MLE) que de son étude au cours d'une cinétique de stress, il n'est pas aisé d'interpréter les résultats obtenus. De plus, concernant les éléments transposables, les paramètres observés sont principalement la transposition et la transposase elle-même, rarement l'expression de son gène. En effet, nos résultats sont en accord avec ceux de Lampe et al. (1998) qui ont mis en évidence une activation de la transposition du MLE *Himar1* de la mouche *Haematobia irritans* à haute température. Au contraire, il a été démontré chez le muflier *Antirrhinum majus*, que la transposition de l'élément *Tam3* n'avait pas lieu lors d'une exposition à une température élevée (25 °C) (Hashida et al., 2003). Kitamura, en 2001, a imputé cette inhibition à l'augmentation du taux de méthylation de la séquence de *Tam3* lors d'une exposition à 25 °C. Mais cette hypothèse a été réfutée en 2006 (Hashida et al., 2006).

Leretour du niveau d'expression du gène de la transposase à un niveau témoin serait dû au fait que chez les MLE, selon une hypothèse, à partir d'un certain niveau de concentration de transposase, la transposition est inhibée (Hartl et al., 1997 ; Lohe et Hartl, 1996). Il peut être envisagé un mécanisme semblable au niveau transcriptionnel.

Dans le cas d'une exposition à une température de 16 °C, une sur-expression du gène de la transposase est observée après 2 h de stress et, comme évoqué pour un stress de 8 h à 32 °C, l'amplitude de variation de l'expression après 30 min d'exposition pourrait s'expliquer par l'intervention d'un deuxième manipulateur. On peut donc penser qu'une sur-expression a lieu dès 30 min et jusqu'à 2 h d'exposition. Ces résultats sont corroborés par les travaux de Hashida (2003) et Kitamura (2001) qui ont mis en évidence une expression de la transposase chez le mufler *Antirrhinum majus* lors d'une exposition à une basse température (15 °C).

En mettant en relation les résultats obtenus pour ces deux gènes (hsp70 et transposase des MLE) il ressort que le gène de la transposase des MLE s'exprime avant celui de la hsp70 lors d'une exposition à une température de 16 °C et inversement dans le cas d'une température de 32 °C. Ce résultat reste à confirmer chez d'autres diatomées ou microalgues ou encore lors de cinétiques de stress différents. De plus, refaire un réplicat lorsqu'une valeur extrême est observée permettrait de le confirmer.

#### **4.2. Effet d'un excès de zinc**

La concentration de 20 µM de Zn ajouté au milieu de culture avait été choisie par Nguyen-Deroche (2012) pour réduire la croissance de diatomées cultivées pendant plusieurs jours en présence de ce métal. Cette concentration est probablement insuffisante pour affecter notablement l'activité photosynthétique de *A.acutiuscula* pendant de courtes durées (ici 8 h). La faible diminution de l'IP et des rendements quantiques après 30 min de contact pourrait être imputable à l'effet osmotique du Zn ajouté. Toutefois, la présence d'exopolysaccharides a probablement limité l'absorption du Zn par les microalgues.

Une concentration plus élevée (40 µM) en Zn ayant un effet sur la photosynthèse de la microalgue et étant plus proche de celles utilisées dans la littérature sur de courtes durées (Nguyen-Derche, 2009) a donc été choisie. Esposito et al. (2012) ont observé une induction de hsp70 dans la mousse aquatique *Leptodictylum riparium* lors de stress métalliques et notamment du zinc. Les auteurs suggèrent que les hsp70 seraient



impliquées dans la résistance aux effets toxiques des métaux et notamment du zinc. De la même façon, Wong et al. (2010) ont montré que des hsp70 de diatomées avaient une expression en protéines accrues lors d'un traitement à l'oxyde de zinc et /ou aux nanoparticules d'oxyde de zinc. Parallèlement, aucune sur-expression ou sous-expression n'a été clairement mise en évidence pour les MLE. Une tendance à la surexpression est visible pour les temps 30 min, 2 h et 8 h. Après 5 h d'exposition au Zn, la moyenne est décalée vers une sous expression. Il faudrait poursuivre cette étude afin de déterminer quel est le facteur déclenchant l'expression des gènes : soit une entrée massive de Zn dans les cellules soit le choc osmotique provoqué par l'ajout de  $\text{ZnCl}_2$  dans le milieu de culture.

L'ensemble de nos résultats montre que la diatomée *A. acutiuscula* réagit très rapidement aux modifications de son environnement, des réponses au niveau physiologique et moléculaire sont observables dès 30 min de traitement (Tab. 19). La réponse moléculaire est généralement plus tardive (observée après 2 h de traitement) que la réponse photosynthétique.

#### **4.3. Effet d'un excès de cuivre**

Chez les diatomées marines, un excès de Cu affecte les paramètres photosynthétiques et l'expression des gènes (Cid et al., 1995 ; Davis et al., 2006, Masmoudi et al., 2013). L'eau de mer enrichie en Cu 3  $\mu\text{M}$  réduit sur le long terme la croissance et les paramètres photosynthétiques de *A. acutiuscula* (Nguyen-Deroche et al. 2009). Pendant les deux premières heures de traitement, l'IP augmente alors que  $\Phi\text{PSII}$  diminue, puis l'IP décroît ainsi que  $\Phi\text{PSII}$ . Ces premiers résultats montrent que l'appareil photosynthétique semble rapidement affecté par l'excès de Cu. Comme précédemment, la production d'exopolysaccharides réduit probablement la toxicité de cet excès de Cu (Croot et al., 2000, Tien et al., 2005, Nguyen-Deroche et al. 2009).

La mauvaise qualité des ARN extraits des microalgues exposées à un excès de Cu ne nous a pas permis de poursuivre l'étude de l'expression des gènes. La mise au point d'un nouveau protocole d'extraction des ARN à partir d'extraits riches en Cu nécessitant trop de temps, les expérimentations avec le Cu ont été abandonnées.

Les deux types de traitement, thermique et métallique, semblent donc entraîner des réponses physiologiques (photosynthèse) et moléculaire différentes. Les hsp70 seraient plus sollicitées lors de fortes concentrations en zinc que lors de variations de température.

## CONCLUSIONS ET PERSPECTIVES

Dans le présent travail, il a été montré la présence de transposons de la famille MLE dans le génome de la diatomée marine *A. acutiuscula*. L'hôte contient des éléments complets *Aacutmar1* ainsi que 3 types d'éléments tronqués *Aacutmar 2, 3 et 4*. Ces éléments appartiennent à une nouvelle sous-famille car ils possèdent un motif catalytique DD43D unique à ce jour dans ce type d'éléments transposables. Il ressort également que plusieurs transposons *Aacutmar* à ORF sont présents de façon naturelle dans le génome d'*A. acutiuscula*, ce qui suggère une activité potentielle de ces éléments. Le nombre d'éléments *Aacutmar1* complets dans le génome, environ 21, suggère également une activité endogène qui, si elle n'est plus effective, l'a été au cours d'une période relativement récente. Parallèlement, l'expression des MLE lorsque les algues sont soumises à des conditions variables notamment de température est également un argument en faveur d'une activité probable des transposons *Aacutmar*.

Il serait maintenant intéressant d'étudier leur capacité à transposer. La mobilité potentielle des transposons pourrait être étudiée parallèlement *in vitro* et *in vivo*. *In vitro*, à l'aide de la construction de vecteurs d'expression et de la production de transposase par des cellules en culture. Ainsi, les différentes étapes de la transposition, capacité de l'enzyme à s'accrocher aux ITR, à couper le transposon, à le réinsérer dans un ADN cible, pourraient être étudiées séparément. *In vivo*, la transposition pourrait être induite dans les microalgues par des chocs thermiques, puis le déplacement des MLE suivi sur les ADN extraits et ce, par des méthodes telles que la SSAP (sequence specific amplification polymorphism) ou encore le Southern blot, deux méthodes qui permettent de localiser les transposons ayant changé de place dans le génome. Si la capacité de transposition des MLE *Aacutmar* était avérée, leur utilisation comme outil de mutagenèse et de transgénèse pourrait alors être étudiée et servir notamment à la transformation d'autres microalgues.

Ce nouvel élément *Aacutmar* constitue très probablement un point de départ à une étude plus exhaustive de la richesse en éléments transposables de la diatomée marine *A. acutiuscula*.



# BIBLIOGRAPHIE

## A

- Abid O, Sellami-Kammoun A, Ayadi H, Drira Z, Bouain A, Aleya L** (2008) Biochemical adaptation of phytoplankton to salinity and nutrient gradients in a coastal solar saltern, Tunisia. *Estuarine Coastal and Shelf Sci* **80**: 391-400
- Allakhverdiev S, Kreslavski V, Klimov V, Los D, Carpentier R, Mohanty P** (2008). Heat stress: an overview of molecular responses in photosynthesis. *Photosynth Res* **98**: 541-550
- Arkhipova IR, Meselson M** (2005) Diverse DNA transposons in rotifers of the class Bdelloidea. *Proc Natl Acad Sci USA* **102**: 11781-1178
- Armbrust EV** (2009) The life of diatoms in the world's oceans. *Nature* **459**: 185-192
- Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WW, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS** (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306**: 79-86
- Augé-Gouillou C, Brillet B, Germon S, Hamelin MH, Bigot Y** (2005) *Mariner Mos1* transposase dimerizes prior to ITR binding. *J Mol Biol* **351**: 117-130
- Augé-Gouillou C, Hamelin MH, Demattei MV, Periquet G, Bigot Y** (2001a) The ITR binding domain of the *Mariner Mos-1* transposase. *Mol Genet Genomics* **265**: 58-65
- Augé-Gouillou C, Hamelin MH, Demattei MV, Periquet M, Bigot Y** (2001b) The wild-type conformation of the *Mos-1* inverted terminal repeats is suboptimal for transposition in bacteria. *Mol Genet Genomics* **265**: 51-57

## B

- Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD** (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO<sub>2</sub>-concentrating mechanisms in algae. *Can J Bot* **76**: 1052-1071

- Bailleul B, Rogato A, de Martino A, Coesel S, Cardol P, Bowler C, Falciatore A, Finazzi G** (2010) An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proc Natl Acad Sci USA* **107**: 18214-18219
- Balciunas D, WangensETen KJ, Wilber A, Bell J, Geurts A, Sivasubbu S, Wang X** (2006) Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genetics* **2**: e169
- Baldauf SL** (2008) An overview of the phylogeny and diversity of eukaryotes. *J Systematics and Evol* **46**: 263-273
- Ballester M, Castelló A, Ibáñez E, Sánchez A, Folch JM** (2004) Real-time quantitative PCR-based system for determining transgene copy number in transgenic animals. *Biotechniques* **37**: 610-613
- Banks JA, Masson P, Fedoroff N** (1988) Molecular mechanisms in the developmental regulation of the maize suppressor-mutator transposable element. *Genes Dev* **2**: 1364-1380
- Banzet N, Richaud C, Deveau Y, Kazmaier M, Gagnon J, Triantaphylides C** (1998) Accumulation of small heat shock proteins, including mitochondrial HSP22, induced by oxidative stress and adaptative response in tomato cells. *The Plant J* **13**: 519-527
- Barry EG, WitherspoonDJ, Lampe DJ** (2004) A bacterial genetic screen identifies functional coding sequences of the insect *mariner* transposable element *Famar1* amplified from the genome of the earwig, *Forficula auricularia*. *Genetics* **166**: 823-833
- Becker B, Hoef-Emden K, Melkonian M** (2008) Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. *BMC Evolutionary Biology* **8**: 203
- Beguiristain T, Grandbastien MA, Puigdomenech P, Casacuberta JM** (2001) Three *Tnt1* subfamilies show different stress-associated patterns of expression in Tobacco. Consequences for retrotransposon control and evolution in plants. *Plant Physiol* **127**: 212-221
- Bertrand M, Poirier I** (2005) Photosynthetic organisms and excess of metals. *Photosynthetica* **43**: 345-353
- Bayer-Giraldi M, Weikusat I, Besir H, Dieckmann G** (2011) Characterization of an antifreeze protein from the polar diatom *Fragilariopsis cylindrus* and its relevance in sea ice. *Cryobiology* **63**: 210-219
- Bigot Y, Brillet B, Augé-Gouillou C** (2005) Conservation of Palindromic and Mirror Motifs within Inverted Terminal Repeats of *mariner*-like Elements. *J Mol Biol* **351**: 108-116

- Biémont C, Vieira C** (2006) Genetics: Junk DNA as an evolutionary force. *Nature* **443**: 521-524
- Bijlsma R, Loeschcke V** (1997) *Environmental Stress, Adaptation and Evolution*. Basel, Switzerland: Birkhauser Verlag. Bjedov, I
- Bischerour J, Lu C, Roth DB, Chalmers R** (2009) Base flipping in V(D)J recombination: insights into the mechanism of hairpin formation and the 12/23 rule. *Mol Cell Biol* **29**: 5889-5899
- Blackman RK, Grimaila R, Koehler MMD, Gelbart WM** (1987) Mobilisation of *hobo* elements residing within the decapentaplegic gene complex: suggesting of a new hybrid dysgenesis system in *Drosophila melanogaster*. *Cell* **49**: 497-505
- Bonin DJ, Droop MR, Maestrini SY, Bonin MC** (1986) Physiological features of six microalgae to be used as indicators of seawater quality. *Crypt Algal* **7**: 23-83
- Bordenstein SR, Reznikoff WS** (2005) Mobile DNA in obligate intracellular bacteria. *Nature Reviews Microbiology* **3**: 688-699
- Boston RS, Viitanen PV, Vierling E** (1996) Molecular chaperones and protein folding in plants. *Plant Mol Biol* **32**: 191-222
- Bourque G** (2009) Transposable elements in gene regulation and in the evolution of vertebrate genomes. *Cur Opin Genet Dev* **19**: 607-612
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otiillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kröger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jézéquel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siaut M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van de Peer Y, Grigoriev IV** (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* **456**: 239-244
- Brillet B, Bigot Y, Augé-Gouillou C** (2007) Assembly of the *Tc1* and *mariner* transposition initiation complexes depends on the origins of their transposase DNA binding domains. *Genetica* **130**: 105-120

- Brunet F, Godin F, Bazin C, David JR, Capy P** (1996) The *mariner* transposable element in natural populations of *Drosophila teissieri*. *J Mol Evol* **42**: 669-675
- Bui QT, Casse N, Leignel V, Nicolas V, Chénais B** (2008) Widespread occurrence of *mariner* transposons in coastal crabs. *Mol Phylogenet Evol* **47**: 1181-1189
- Bui QT, Delaurière L, Casse N, Nicolas V, Laulier M, Chénais B** (2007) Molecular characterization and phylogenetic position of a new *mariner*-like element in the coastal crab, *Pachygrapsus marmoratus*. *Gene* **396**: 248-256
- Bureau ET, Wessler SR** (1992) *Tourist*: a large family of small inverted repeat elements frequently associated with maize genes. *Plant Cell* **4**: 1283-1294
- Bureau ET, Wessler SR** (1994) *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. *Plant Cell* **6**: 907-916
- Burki F, Shalchian-Tabrizi K, Minge M, Skjæveland A, Nikolaev SI, Jakobsen KS, Pawlowski J** (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PLoS One* **2**: e790
- Burki F, Shalchian-Tabrizi K, Pawlowski J** (2008) Phylogenomics reveals a new « megagroup » including most photosynthetic eukaryotes. *Biol Lett* **4**: 366-369

## C

- Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, Jacobsen SE** (2003) Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol* **13**: 2212-2217
- Caprioara-Buda M, Meyer W, Jeynov B, Corbisier P, Trapmann S, Emons H** (2012) Evaluation of plasmid and genomic DNA calibrants used for the quantification of genetically modified organisms. *Anal Bioanal Chem* **404**: 29-42
- Capy P, Maruyama K, David JR, Hartl DL** (1991) Insertion sites of the transposable element *mariner* are fixed in the genome of *Drosophila sechellia*. *J Mol Evol* **33**: 450-456
- Capy P** (2004) Structure et évolution des éléments transposables. *Journal de la Société de Biologie* **198**: 393-398
- Capy P, Anxolabéhère D, Langin T** (1994) The strange phylogenies of transposable elements: are horizontal transfers the only explanation ? *Trends Genet* **10**: 7-12

- Capy P, Bazin C, Higuët D, Langin T** (1998) Dynamics and evolution of transposable elements. Springer, Austin
- Capy P, Gasperi G, Biémont C, Bazin C** (2000) Stress and transposable elements : co-evolution or useful parasites? *Heredity* **85**: 101-106
- Carpenter R, Martin C, Coen E** (1987) Comparison of genetic behaviour of the transposable element Tam3 at two unlinked pigment loci in *Antirrhinum majus*. *Mol Gen Genet* **207**: 82-89
- Carpentier G, Jaillet J, Pflieger A, Adet J, Renault S, Augé-Gouillou C** (2011) Transposase-transposase interactions in MOS1 complexes: a biochemical approach. *J Mol Biol.* **405**: 892-908
- Casse N, Bui, QT, Nicolas V, Renault S, Bigot Y, Laulier M** (2006) Species sympatry and horizontal transfers of *mariner* transposons in marine crustacean genomes. *Mol Phylogen Evol* **40**: 609-619
- Cavalier-Smith T** (1999). Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, Dinoflagellate, and Sporozoan plastid origins and the Eukaryote family tree. *J Eukaryot Microbiol* **46**: 347-366
- Cavalier-Smith T** (2010) Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree. *Biol Lett* **6**: 342-345
- Chandler VL, Walbot V** (1986) DNA modification of a maize transposable element correlates with loss of activity. *Proc Natl Acad Sci USA* **83**: 1767-1771
- Chen YC** (2012) The biomass and total lipid content and composition of twelve species of marine diatoms cultured under various environments. *Food Chem* **131**: 211-219
- Chepurnov VA, Mann DG, Sabbe K, Vyverman W** (2004) Experimental studies on sexual reproduction in diatoms. *Int Rev Cytol* **237**: 91-154.
- Chénais B, Caruso A, Hiard S, Casse N** (2012) The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. *Gene* **1**: 7-15
- Chisti Y** (2007) Biodiesel from algae. *Biotechnol Adv* **25**: 294-306
- Cid A, Herrero C, Torres E, Abalde J** (1995) Copper toxicity on the marine microalga *Phaeodactylum tricornutum*: effect on photosynthesis and related parameters. *Aquatic Toxicol* **31**: 165-174
- Cooksey KE** (1981) Requirement for calcium in adhesion of a fouling diatom to glass. *Appl. Environ. Microbiol* **41**: 1378-1382



**Cordeaux R** (2009) Gene conversion maintains nonfunctional transposable elements in an obligate mutualistic endosymbiont. *Mol Biol Evol***26**: 1679-1682

**Croot PL, Moffett JW, Brand LE** (2000) Production of extracellular Cu complexing ligands by eucaryotic phytoplankton in response to Cu stress. *Limnol Oceanogr* **45**: 619-627

## **D**

**Damsté JSS, Muyzer G, Abbas B, Rampen SW, Massé G, Allard WG, Belt ST, Robert JM, Rowland SJ, Moldowan JM, Barbanti SM, Fago FJ, Denisevich P, Dahl J, Trindade LAF, Schouten S** (2004) The rise of the Rhizosolenid diatoms. *Science***304**: 584-587

**Danilov, R, Ekelund N** (2001). Effects of Cu<sup>2+</sup>, Ni<sup>2+</sup>, Pb<sup>2+</sup>, Zn<sup>2+</sup> and pentachlorophenolon photosynthesis and motility in *Chlamydomonas reinhardtii* in short-term exposure experiments. *BMC Ecol* **1**:1

**Day A, Schirmer-Rahire M, Kuchka MR, Mayfield SP, Rochaix JD** (1988) A transposon with unusual arrangement of long terminal repeats in the green alga *Chlamydomonas reinhardtii*. *EMBO J* **7**: 1917-1927

**Davis AK, Hildebrand M, Palenik B** (2006) Gene expression induced by copper stress in the diatom *Thalassiosira pseudonana*. *Eukaryot Cell***5**: 1157-1168

**Dawson A, Finnegan DJ** (2003) Excision of the Drosophila mariner transposon *Mos1*. Comparison with bacterial transposition and V(D)J recombination. *Mol Cell* **11**: 225-235.

**Delaunay L, Chénais B, Pradier E, Hardivillier Y, Renault S, Casse N** (2009) DNA binding specificity and cleavage activity of Pacman transposase. *Biochemistry***48**: 7279-7286

**Delieu T, Walker DA** (1972) An improved cathode for the measurement of photosynthetic oxygen evolution by isolated chloroplasts. *New Phytologist***71**: 201-225

**Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, Partensky F, Degroeve S, Echeynie S, Cooke R, Saeys Y, Wuyts J, Jabbari K, Bowler C, Panaud O, Piegou B, Ball S, Ral JP, Bouget FY, Piganeau G, De Baets B, Picard A, Delseny M, Demaille J, Van de Peer Y, Moreau H** (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci USA* **103**:11647-11652

**Deschamps P, Moreira D** (2012) Reevaluating the green contribution to diatom genomes. *Genome Biol Evol***4**: 683-688

- Devos KM, Brown JK, Bennetzen JL** (2002) Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Res* **12**: 1075-1079
- Diao X, Freeling M, Lisch D** (2006) Horizontal transfer of a plant transposon. *PLoS Biology* **4**: e5
- Diao XM, Lisch D** (2006) *Mutator* transposon in maize and MULEs in plant genome. *Acta Genet Sin* **33**: 477-487
- Dinesh-Kumar SP, Whitham S, Choi D, Hehl R, Corr C, Baker B** (1995) Transposon tagging of tobacco mosaic virus resistance gene N<sup>+</sup>: Its possible role in the TMV-N-mediated signal transduction pathway. *Proc Natl Acad Sci USA* **92**: 4175-4180
- Ding Y, Wang X, Su L, Zhai J, Cao S, Zhang D, Liu C, Bi Y, Qian Q, Cheng Z, Chu C, Cao X** (2007) SDG714, a histone H3K9 methyltransferase, is involved in *Tos17* DNA methylation and transposition in rice. *Plant Cell* **19**: 9-22
- Dombroski B, Mathias S, Nanthakumar E, Scott A, Kazazian H** (1991) Isolation of an active human transposable element. *Science* **254**: 1805-1808
- Domergue F, Lerchl J, Zähringer U, Heinz E** (2002) Cloning and functional characterization of *Phaeodactylum tricornutum* front-end desaturases involved in eicosapentanoic acid biosynthesis. *Eur J Biochem* **269**: 4105-4113
- Dongsansuk A, Lütz C, Neuner G** (2013) Effects of temperature and irradiance on quantum yield of PSII photochemistry and xanthophyll cycle in a tropical and a temperate species *Photosynthetica* **51**: 13-21
- Doyle JJ, Doyle JL** (1990) Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15
- Dyerberg J, Jorgensen KA** (1982) Marine oils and thrombogenesis. *Progress in Lipid Research* **21**: 255-269

## E

- Eagle SHC, Crease TJ** (2012) Copy number variation of ribosomal DNA and Pokey transposons in natural populations of *Daphnia*. *Mob DNA* **3**: 4
- Ellis THN, Poyser SJ, Knox MR, Vershini AV, Ambrose MJ** (1998) Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. *Mol Gen Genet* **260**: 9-19

**Esposito S, Sorbo S, Conte B, Basile A** (2012) Effects of heavy metals on ultrastructure and HSP70s induction in the aquatic moss *Leptodictyum riparium* Hedw. *Int J Phytoremediation* **14**: 443-455

**Evans KM, Wortley AH, Mann DG** (2007) An assessment of potential diatom “Barcode” genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* **158**: 349-364

## **F**

**Falkowski PG, Barber RT, Smetacek V** (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**: 200-206

**Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR** (2004) The evolution of modern eucaryotic phytoplankton. *Science* **305**: 354-360

**Falkowski PG, Raven JA** (1997) *Aquatic photosynthesis. Blackwater Science*, London

**Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain SH, Strauss SH, Halpern ME, Ukomadu C, Sadler KC, Pradhan S, Pellegrini M, Jacobsen SE** (2010) Conservation and divergence of methylation patterning in plants and animals. *Proc Nat Acad Sci USA* **107**: 8689-8694

**Ferris PJ** (1989) Characterization of a *Chlamydomonas* transposon, Gulliver, resembling those in higher plants. *Genetics* **122**: 363-377

**Feschotte C** (2008) Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* **9**: 397-405

**Feschotte C, Pritham EJ** (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* **41**: 331-368

**Feschotte C, Wessler SR** (2002) Mariner-like transposases are widespread and diverse in flowering plants. *Proc Natl Acad Sci USA* **99**: 280-285

**Feschotte C, Swamy L, Wessler SR** (2003) Genome-wide analysis of mariner-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITEs). *Genetics* **163**: 747-758

**Feschotte C, Jiang N, Wessler SR** (2002) Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* **3**: 32-341

**Field CB, Behrenfeld MJ, Danderson JT, Falkowski P** (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237-240

**Finnegan DJ** (1989) Eukaryotic transposable elements and genome evolution. Trends in Genetics **5**: 103-107

**Forreiter C, Kirschner M, Nover L** (1997) Stable transformation of an Arabidopsis cell suspension culture with firefly luciferase providing a cellular system for analysis of chaperone activity in vivo. Plant Cell **9**: 2171-2181

**Fortune P, Roulin A, Panaud O** (2008) Horizontal transfer of transposable elements in plants. Communicative and integrative biology **1**: 1- 4

**Franz G, Savakis C** (1991) *Minos*, a new transposable element from *Drosophila hydei*, is a member of the *Tc1*-like family of transposons. Nucleic Acids Res **19**:6646

## G

**Garcia-Fernàndez J, Marfany G, Baguñà J, Saló E** (1993) Infiltration of mariner elements. Nature **364**: 109-110

**Gaspar T, Franck T, Bisbis B, Kevers C, Jouve L, Hausman JF, Dommes J** (2002) Concepts in plant stress physiology. Application to plant tissue cultures. Plant Growth Regul **37**:263-285

**Gilbert C, Schaack S, Pace II JK, Brindley PJ, Feschotte C** (2010) A role for host-parasite interactions in the horizontal transfer of transposons across phyla. Nature **464**: 1347-1350

**Giordano M, Beardall J, Raven JA** (2005) CO<sub>2</sub> concentrating mechanisms in algae. Mechanisms, environmental modulation and evolution. Ann Rev Plant Biol **56**: 99-131

**Giraud T, Capy P** (1996) Somatic activity of the mariner transposable element in natural populations of *Drosophila simulans*. Proc Biol Sci **263**: 1481-1486

**Gomulski LM, Torti C, Malacrida AR, Gasperi G** (1997) Ccmar1, a full-length mariner element from the Mediterranean fruit fly, *Ceratitis capitata*. Insect Mol Biol **6**:241-253

**Gomulski LM, Torti C, Bonizzoni M, Moralli D, Raimondi E, Capy P, Gasperi G, Malacrida AR** (2001) A new basal subfamily of mariner elements in *Ceratitidis rosa* and other tephritid flies. J Mol Evol **53**:597-606

**González J, Petrov D** (2009) MITEs - the ultimate parasites. Science **325**: 1352-1353

**Gould SB, Waller RF, McFadden GI** (2008) Plastidevolution. Annu Rev Plant Biol **59**: 491-517

**Govindjee** (1995) Sixty-three years since Kautsky : chlorophyll *a* fluorescence. Aust J Plant Physiol **22**: 131-160

- Grandbastien MA** (1998) Activation of plant retrotransposons under stress conditions. *Trends Plant Sci* **3**: 181-187
- Grandbastien MA, Audéon C, Bonnivard E, Casacuberta JM, Chalhoub B, Costa AP, Le QH, Melayah D, Petit M, Poncet C, Tam SM, Van Sluys MA, Mhiri C** (2005) Stress activation and genomic impact of *Tnt1* retrotransposons in Solanaceae. *Cytogenet. Genome Res* **110**: 229-241
- Graham JE, Spanier JG, Jarvik JW** (1995) Isolation and characterization of *Pioneer1*, a novel *Chlamydomonas* transposable element. *Curr Genet* **28**: 429-436
- Granum E, Kikvold S, Myklestad SM** (2002) Cellular and extracellular production of carbohydrates and amino acids by marine diatom *Skeletonema costatum*: diel variations and effects of N depletion. *Mar Ecol Prog Ser* **242**: 82-93
- Grossman GL, Cornel AJ, Rafferty CS, Robertson HM, Collins FH** (1999) Tsessebe, Topi and Tiang: three distinct Tc1-like transposable elements in the malaria vector, *Anopheles gambiae*. *Genetica* **105**: 69-80
- Guihéneuf F, Fouqueray M, Mimouni V, Ulmann L, Jacquette B, Tremblin G** (2010). Effect of UV stress on the fatty acid and lipid class composition in two marine microalgae: *Pavlovalutheri* (Pavlovophyceae) and *Odontella aurita* (Bacillariophyceae). *J Appl Phycol* **22**: 629-638
- Guihéneuf F, Mimouni V, Ulmann L, Tremblin G** (2008) Environmental factors affecting growth and oméga 3 fatty acid composition in *Skeletonema costatum*. The influences of irradiance and carbon source. *Diatom Research* **23**: 93-103
- Guy CL, Li QB** (1998) The organization and evolution of the spinach stress 70 molecular chaperone gene family. *Plant Cell* **10**: 539-556

## H

- Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, Bhattacharya D** (2007) Phylogenomic analysis supports the monophyly of Cryptophytes and Haptophytes and the association of Rhizaria with Chromalveolates. *Mol Biol Evol* **24**: 1702-1713
- Hackett PB, Ekker SC, Largaespada DA, McIvor RS** (2004) *Sleeping Beauty* transposon-mediated gene therapy for prolonged expression. *Adv Genet* **54**: 189-232
- Hall TA** (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98

- Han Y, Wessler SR** (2010) MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res* **38**: e199
- Han Y, Qin S, Wessler SR** (2013) Comparison of class2 transposable elements at superfamily resolution reveals conserved and distinct features in cereal grass genomes. *BMC Genomics* **14**: 71
- Hanahan D, Bloom FR** (1996) Mechanisms of DNA transformation. F.C. Neidhardt (Ed.), *Escherichia coli* and *Salmonella*- cellular and molecular biology (2nd Ed.). American Society for Microbiology, Washington, D.C: 2449-2459
- Harper JT, Waanders E, Keeling PJ** (2005) On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes. *Int J Syst Evol Microbiol* **55**: 487-496
- Harrison BJ, Fincham JRS** (1964) Instability at the pal locus in *Antirrhinum majus*: I. Effects of environment on frequencies of somatic and germinal mutation. *Heredity* **19**: 237-258
- Harrison PJ, Waters RE, Taylor FJR** (1980) A broad spectrum artificial seawater medium for coastal and open ocean phytoplankton. *J Phycol* **16**: 28-35
- Hartl DL** (2001) Discovery of the transposable element *mariner*. *Genetics* **157**: 471-476
- Hartl DL, Lohe AR, Lozovskaya ER** (1997a) Modern thoughts on an ancient *mariner*: function, evolution, regulation. *Annu Rev Genet* **31**: 337-358
- Hartl DL, Lozovskaya ER, Nurminsky DI, Lohe AR** (1997b) What restricts the activity of *mariner*-like transposable elements ? *Trends Genet* **13**: 197-201
- Harwood DM, Gersonde R** (1990) Lower Cretaceous diatom from ODP Leg 113 site 693 (Weldell Sea), part 2, resting spores, Chrysophycean cysts, and endoskeletal dinoflagellates and notes on the origin of diatoms. In. *Proceeding Ocean Drilling Program of Sciences Results*. P. F. Barker and J. P. Kennett (Eds), pp 403-426
- Hashida SN, Kitamura K, Mikami T, Kishima Y** (2003) Temperature shift coordinately changes the activity and the methylation state of transposon *Tam3* in *Antirrhinum majus*. *Plant Physiology* **132**: 1207-1216
- Hashida SN, Uchiyama T, Martin C, Kishima Y, Sano Y, Mikami T** (2006) The temperature-dependent change in methylation of the *Antirrhinum* transposon *Tam3* is controlled by the activity of its transposase. *Plant Cell* **18**: 104-118
- Hawkins JS, Kim H, Nason JD, Wing RA, Wendel JF** (2006) Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*. *Genome Res* **16**: 1252-1261

- Hermann D** (2011) Thèse de doctorat : Caractérisation d'éléments transposables de type *mariner* chez les microalgues marines. Université du Maine, 261 pages
- Hildebrand M, Corey DK, Ludwig JR, Kukel A, Feng TY, Volcani BE** (1991) Plasmids in diatom species. *J Bacteriol* **173**: 5924-5927
- Hiom K, Melek M, Gellert M** (1998) DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* **94**: 463-470
- Hoffmann AA, Parsons PA** (1997) *Extreme Environmental Change and Evolution*. Cambridge, England: Cambridge University Press.
- Holyoake AJ, Kidwell MG** (2003) *Vege* and *Mar*: two novel hAT MITE families from *Drosophila willistoni*. *Mol Biol Evol* **20**: 163-167
- Houck MA, Clark JB, Peterson KR, Kidwell MG** (1991) Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* **253**: 1125-1128

## I

- Ifremer** (2006). Guide d'information. Phycotoxines amnésiantes.
- Ishikawa C, Tafuku S, Kadekaru T, Sawada S, Tomita M, Okudaira T, Nakazato T, Toda T, Uchihara JN, Taira N, Ohshiro K, Yasumoto T, Ohta T, Mori N** (2008) Antiadult T-cell leukemia effects of brown algae fucoxanthin and its deacetylated product, fucoxanthinol. *Int J Cancer* **123**: 2702-2712
- Ivics Z, Hackett PB, Plasterk RH, Izsvák Z** (1997) Molecular reconstruction of *Sleeping Beauty*, a *Tc1*-like transposon from fish, and its transposition in human cells. *Cell* **91**: 501-510
- Ivics Z, Izsvák Z** (2006) Transposons for gene therapy. *Current gene therapy* **6**: 593-607
- Ivics Z, Izsvak Z, Minter A, Hackett PB** (1996) Identification of functional domains and evolution of *Tc1*-like transposable elements. *Proc Natl Acad Sci USA* **93**: 5008-5013
- Izsvak Z, Ivics Z** (2004) *Sleeping Beauty* Transposition: Biology and Applications for Molecular Therapy. *Mol Therapy* **9**: 147-156

## J

- Jacobs G, Dechyeva D, Menzel G, Dombrowski C, Schmidt T** (2004) Molecular characterization of *Vulmar1*, a complete *mariner* transposon of sugar beet and diversity of *mariner*- and *En/Spm*-like sequences in the genus *Beta*. *Genome* **47**: 1192-1201
- Jacobson JW, Medhora MM, Hartl DL**(1986) Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc Natl Acad Sci USA* **83**: 8684-8688
- Jaillet J, Genty M, Cambefort J, Rouault JD, Augé-Gouillou C** (2012) Regulation of *mariner* transposition: the peculiar case of *Mos1*. *PLoS One* **7**: e43365
- Jain R, Rivera MC, Moore JE, Lake JA** (2002) Horizontal gene transfer in microbial genome evolution. *Theor Popul Biol* **61**: 489-495
- Jarvik T, Lark KG** (1998) Characterization of *Soymar1*, a *mariner* element in soybean. *Genetics* **149**: 1569-1574
- Jehle JA, Nickel A, Vlak JM, Backhaus H** (1998) Horizontal escape of the novel Tc1-like lepidopteran transposon TCp3.2 into *Cydia pomonella* granulovirus. *J Mol Evol* **46**: 215-224
- Jeyaparakash A, Hoy MA** (1995) Complete sequence of a *mariner* transposable element from the predatory mite *Metaseiulus occidentalis* isolated by an inverse PCR approach. *Insect Mol Biol* **4**: 31-39
- Jiang N, Feschotte C, Zhang X, Wessler SR** (2004) Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). *Curr Opin Plant Biol* **7**: 115-119
- Jiang H, Gao K** (2004) Effects of lowering temperature during culture on the production of polyunsaturated fatty acids in the marine diatom *Phaeodactylum tricornutum* (Bacillariophyceae). *J Phycol* **40**: 651-654
- Jiang N, Bao Z, Zhang X, Hirichika H, Eddy SR, McCouch SR, Wessler SR** (2003) An active DNA transposon family in rice. *Nature* **421**: 163-166
- Jupin H, Lamant A** (1997) *La photosynthèse*. Ed. Masson, Paris



## K

- Kaczmarska I, Poulickova A, Sato S, Edlund MB, Idei M, Watanabe T, Mann DG** (2013) Proposals for a terminology for diatom sexual reproduction, auxospores and resting stages. *Diatom Research*, doi:10.1080/0269249X.2013.791344
- Kapitonov VV, Jurka J** (2005) RAG1 core and V(D)J recombination signal sequences were derived from *Transib* transposons. *PLoS Biology***3**: e181
- Kapitonov VV, Jurka J** (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet***9**: 411-412
- Kass DH, Jamison N, Mayberry MM, Tecle E** (2007) Identification of a unique Alu-based polymorphism and its use in human population studies. *Gene***390**: 146-152
- Keeling PJ, Archibald JM, Fast NM, Palmer JD** (2004) Comment on "The evolution of modern eucaryotic phytoplankton". *Science* **306**: 2191
- Kidwell MG, Lisch DR** (2001) Perspectives: transposable elements, parasitic DNA, and genome. *Evolution* **55**: 1-24
- Kidwell MG** (1977) Reciprocal differences in female recombination associated with hybrid dysgenesis in *Drosophila melanogaster*. *Genet Res* **30**: 77-88
- Kidwell MG, Lish D** (1997) Transposable elements as sources of variation in animals and plants. *Proc Natl Acad Sci USA* **94**: 7704-7711
- Kitamura K, Hashida S, Mikami T, Kishima Y** (2001) Position effect of the excision frequency of the *Antirrhinum* transposon *Tam3* : implications for the degree of position-dependent methylation in the ends of the element. *Plant Mol Biol* **47**: 475-490
- Kolacsek O, Krízsik V, Schamberger A, Erdei Z, Apáti Á, Várady G, Mátés L, Izsvák Z, Ivics Z, Sarkadi B, Orbán T** (2011) Reliable transgene-independent method for determining Sleeping Beauty transposon copy numbers. *Mob DNA* **2**: 5
- Kosugi S, Hasebe M, Tomita M, Yanagawa H** (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci USA***106**:10171-10176
- Kröger N, Sumper M** (1998) Diatom cell wall proteins and the cell biology of silica biomineralization. *Protist***149**: 213-219
- Kutschera U, Niklas KJ** (2005) Endosymbiosis, cell evolution, and speciation. *Theory Biosci* **124**: 1-24

**Külheim C, Agren J, Jansson S** (2002) Rapid régulation of light-harvesting and plant fitness in the field. *Science* **297**: 91-93

## **L**

**Lagarde M, Croset M, Sicard B, Dechavanne N** (1986) Biological activities and metabolism of eicosenoic acids in relation to platelet and endothelial function. *Progress in Lipid Research* **25**: 269-271

**Laha T, Loukas A, Wattanasatitarpa S, Somprakhon J, Kewgrai N, Sithithaworn P, Kaewkes S, Mitreva M, Brindley PJ** (2007) The bandit, a new DNA transposon from a hookworm - Possible horizontal genetic transfer between host and parasite. *PLoS Negl Trop Dis* **1**: e35

**Lampe DJ, Churchill ME, Robertson HM** (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J* **15**: 5470-5479

**Lampe DJ, Walden KKO, Robertson HM** (2001) Loss of transposase-DNA interaction may underlie the divergence of mariner family transposable elements and the ability of more than one *mariner* to occupy the same genome. *Mol Biol Evol* **18**: 954-961

**Lampe DJ, Witherspoon DJ, Soto-Adames FN, Robertson HM** (2003) Recent horizontal transfer of *mellifera* subfamily *mariner* transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. *Mol Biol Evol* **20**: 554-562

**Langin T, Capy P, Daboussi MJ** (1995). The transposable element *impala*, a fungal member of the *Tc1-mariner* superfamily. *Mol Gen Genet* **246**: 19-28

**Lane CE, Archibald JM** (2008) The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol Evol* **23**: 268-275

**Lavaud J, Rousseau B, Etienne AL** (2002a) In diatoms, a transthylakoid proton gradient alone is not sufficient to induce a non-photochemical fluorescence quenching. *FEBS Lett* **523**: 163-166

**Lavaud J, Rousseau B, Etienne AL** (2002b) Photosystem II electron transfer cycle and chlororespiration in planktonic diatoms. *Photosynth Res* **74**: 51-59

**Lavaud J, Rousseau B, van Gorkom HJ, Etienne AL** (2002c) Influence of the diadinoxanthin pool size on photoprotection in the marine planktonic diatom *Phaeodactylum tricornutum*. *Plant Physiol* **129**: 1398-1406

- Lavaud J, Lepetit B** (2013) An explanation for the inter-species variability of the photoprotective non-photochemical chlorophyll fluorescence quenching in diatoms. *Bioch Biophys Acta* **1827**: 294-302
- Le QH, Wright S, Yu Z, Bureau T** (2000) Transposon diversity in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **97**: 7376-7381
- Le Rouzic A, Capy P** (2005) The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. *Genetics* **169**: 1033-1043
- Lee GJ, Vierling E** (2000) A small heat shock protein cooperates with heat shock protein 70 systems to reactivate a heat-denatured protein. *Plant Physiol* **122**: 189-198
- Lerman DN, Michalak P, Helin AB, Bettencourt BR, Feder ME** (2003) Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. *Mol Biol Evol* **20**: 135-44
- Leroy H, Castagnone-Sereno P, Renault S, Augé-Gouillou C, Bigot Y, Abad P** (2003) Characterization of *Mcmar1*, a mariner-like element with large inverted terminal repeats (ITRs) from the phytoparasitic nematode *Meloidogyne chitwoodi*. *Gene* **304**: 35-41
- Levin HL, Moran JV** (2011) Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet* **12**: 615-627
- Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B** (2012) MicroRNA regulation of plant innate immune receptors. *Proc Natl Acad Sci USA* **109**: 1790-1795
- Li TH, Spearow J, Rubin CM, Schmid CW** (1999) Physiological stresses increase mouse short interspersed element (SINE) RNA expression *in vivo*. *Gene* **239**: 367-372
- Lidholm DA, Gudmundsson GH, Boman HG** (1991) A highly repetitive, mariner-like element in the genome of *Hyalophora cecropia*. *J Biol Chem* **266**: 11518-11521
- Lin BL, Wang JS, Liu HC, Chen RW, Meyer Y, Barakat A, Delseny M** (2001) Genomic analysis of the Hsp70 superfamily in *Arabidopsis thaliana*. *Cell Stress Chaperones* **6**: 201-208
- Lippman Z, May B, Yordan C, Singer T, Martienssen R** (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol* **1**: e67
- Liu D, Bischerour J, Siddique A, Buisine N, Bigot Y, Chalmers R** (2007) The human SETMAR protein preserves most of the activities of the ancestral *Hsmar1* transposase. *Mol Cell Biol* **27**: 1125-1132

- Lohe AR, De Aguiar D, Hartl DL** (1997) Mutations in the marinertransposase: the D,D(35)Econsensussequence is nonfunctional. *Proc Natl Acad Sci USA* **94**: 1293-1297
- Lohe AR, Moriyama EN, Lidholm DA, Hartl DL** (1995) Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. *Mol Biol Evol* **12**: 62-72
- Lorenzi H, Thiagarajan M, Haas B, Wortman J, Hall N, Caler E** (2008) Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genomics* **9**:595
- Lukens LN, Zhan S** (2007) The plant genome's methylation status and response to stress: implications for plant improvement. *Curr Opin Plant Biol* **10**: 317-322

## M

- Macas J, Koblízková A, Neumann P** (2005) Characterization of Stowaway MITEs in pea (*Pisum sativum* L.) and identification of their potential master elements. *Genome* **48**: 831-839
- Machida C, Machida Y** (1989) Regulation of *IS1* transposition by the insA gene product. *J Mol Biol* **208**: 567-574
- Mandrioli M** (2003) Identification and chromosomal localization of mariner-like elements in the cabbage moth *Mamestra brassicae* (Lepidoptera). *Chromosome Res* **11**:319-322
- Mann DG** (1999) The species concept in diatoms. *Phycologia* **38**: 437-495
- Martínez G, Slotkin K** (2012) Developmental relaxation of transposable element silencing in plants: functional or byproduct ? *Curr Opin Plant Biol* **15**: 1-7
- Masmoudi S, Nguyen-Deroche N, Caruso A, Ayadi H, Morant-Manceau A, Tremblin G, Bertrand M, Schoefs B** (2013) Cadmium, Copper, Sodium and Zinc Effects on Diatoms: from Heaven to Hell - a Review. *Cryptogamie, Algologie* **34**:185-225
- Mason G, Provero P, Vaira AM, Accotto PA** (2002) Estimating the number of integrations in transformed plants by quantitative real-time PCR. *BMC Biotechnol* **2**:20
- Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, Yoshida Y, Nishimura Y, Nakao S, Kobayashi T, Momoyama Y, Higashiyama T, Minoda A, Sano M, Nomoto H, Oishi K, Hayashi H, Ohta F, Nishizaka S, Haga S, Miura S, Morishita T, Kabeya Y, Terasawa K, Suzuki Y, Ishii Y, Asakawa S, Takano H, Ohta N, Kuroiwa H, Tanaka K, Shimizu N,**

- Sugano S, Sato N, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T** (2004) Genomesequene of the ultrasmallunicellularredalgaCyanidioschyzonmerolae 10D. *Nature***428**: 653-657
- Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ** (2009) RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* **21**: 367-376
- Maumus F, Allen AE, Mhiri C, Hu H, Jabbari K, Vardi A, Grandbastien MA, Bowler C** (2009) Potential impact of stress activated retrotransposons on genome evolution in a marine diatom. *BMC Genomics* **10**: 624
- Maxwell K, Johnson GN** (2000) Chlorophyllfluorescence--a practicalguide. *J Exp Bot* **51**: 659-668
- McClintock B** (1953) Induction of instability at selected loci in maize. *Genetics* **38** (6) : 579-599.
- McClintock B** (1951) Chromosome organization and genic expression. Cold Spring Harbor Symposium on Quantitative Biology**16**: 13-47
- McClintock B** (1984) The significance of responses of the genome to challenge. *Science* **226**:792-801
- Medhora M, Maruyama K, Hartl DL** (1991) Molecular and functional analysis of the mariner mutator element *Mos1* in *Drosophila*. *Genetics* **128**:311-318
- Melayah D, Bonnivard E, Chalhoub B, Audeon C, Grandbastien MA** (2001) The mobility of the tobacco *Tnt1* retrotransposon correlates with its transcriptional activation by fungal factors. *Plant J* **28**: 159-168
- Mhiri C, Morel JB, Vernhettes S, Casacuberta JM, Lucas H, Grandbastien MA** (1997) The promoter of the tobacco *Tnt1* retrotransposon is induced by wounding and abiotic stress. *Plant Mol Biol* **33**: 257-266
- Miller G, Mittler R** (2006) Could heat shock transcription factors function as hydrogen peroxide sensors in plants? *Ann Bot***98**: 279-288
- Mimouni V, Ulmann L, Pasquet V, Mathieu M, Picot L, Bougaran G, Cadoret JP, Morant-Manceau A, Schoefs B** (2012) The potential of microalgae for the productin of bioactive molécules of pharmaceutical interest. *Curr Pharm Biotech* **13**: 2733-2750
- Miskey C, Izsvák Z, Plasterk RH, Ivics Z** (2003) The Frog Prince<sup>®</sup>: a reconstructed transposon from *Rana pipiens* with high transpositional activity in vertebrate cells. *Nucleic Acids Res***31**: 6873-6881

- Miskey C, Izsvak Z, Kawakami K, Ivics Z** (2005) DNA transposons in vertebrate functional genomics. *Cell Mol Life Sci* **62**: 629-641
- Miskey C, Papp B, Mátés L, Sinzelle L, Keller H, Izsvák Z, Ivics Z** (2007) The ancient mariner sails again: transposition of the human *Hsmar1* element by a reconstructed transposase and activities of the SETMAR protein on transposon ends. *Mol Cell Biol* **27**: 4589-4600
- Mittler R** (2002) Oxidative stress, antioxydants and stress tolerance. *Trends in Plant Sci* **7**: 405-410
- Mizuuchi K** (1997) Polynucleotidyl transfer reactions in site-specific DNA recombination. *Genes to Cells* **2**: 1-12
- Mock T, Valentin K** (2004) Photosynthesis and cold acclimation: evidence for a polar diatom. *J. Phycol* **40**: 732-741
- Molino PJ et Wetherbee R** (2008) The biology of biofouling diatoms and their role in the development of microbial slimes. *Biofouling* **24**: 365-379
- Montiel EE, Cabrero J, Camacho JP, López-León MD** (2012). Gypsy, RTE and Mariner transposable elements populate *Eyrepocnemis plorans* genome. *Genetica* **140**: 365-374
- Morant-Manceau A, Nguyen TLN, Pradier E, Tremblin G** (2007) Carbonic anhydrase activity and photosynthesis in marine diatoms. *Eur J Phycol* **42**: 263-273
- Moroney JV, Somanchi A** (1999) How do algae concentrate CO<sub>2</sub> to increase the efficiency of photosynthetic carbon fixation? *Plant Physiol* **119**: 9-16
- Mouget JL, Gastineau R, Davidovich O, Gaudin P, Davidovich NA** (2009) Light is a key factor in triggering sexual reproduction in the pennate diatom *Haslea ostrearia*. *FEMS Microbiol Ecol* **69**: 194-201
- Mouget JL, Tremblin G, Morant-Manceau A, Morancais M, Robert JM** (1999) Long-term acclimation of *Haslea ostrearia* (Bacillariophyta): effect of irradiance on growth rates, pigment content and photosynthesis. *Eur J Phycol* **34**: 109-115
- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D** (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324**: 1724-1726
- Müller-Feuga A** (1997) Microalgues marines : les enjeux de la recherche. Ifremer (Ed)

## N

- Nakajima Y, Fujimoto H, Negishi T, Hashido K, Shiino T, Tsuchida K, Hidaka M, Takada N, Maekawa H** (2002) Possible horizontal transfer of mariner-like sequences into some invertebrates including lepidopteran insects, a grasshopper and a coral. *J. Insect Biotechnol Sericology* **71**:109-121
- Nakamura Y, Sasaki N, Kobayashi M, Ojima N, Yasuike M, Shigenobu Y, Satomi M, Fukuma Y, Shiwaku K, Tsujimoto A, Kobayashi T, Nakayama I, Ito F, Nakajima K, Sano M, Wada T, Kuhara S, Inouye K, Gojobori T, Ikeo K** (2013) The first symbiont-free genome sequence of mariner red alga, *Susabi-nori* (*Pyropia yezoensis*). *PLoS One* **8**: e57122
- Nakatsuka T, Nishihara M, Mishiba K, Hirano H, Yamamura S** (2006) Two different transposable elements inserted in flavonoid 3',5'-hydroxylase gene contribute to pink flower coloration in *Gentiana scabra*. *Mol Genet Genomics* **275**: 231-241
- Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H, Inoue H, Tanisaka T** (2003) Mobilization of a transposon in the rice genome. *Nature* **421**: 170-172
- Nassiri Y, Robert JM, Rincé Y, Ginsburger-Vogel T** (1998) The cytoplasmic fine structure of the diatom *Haslea ostrearia* (Bacillariophyceae) in relation to marennine production. *Phycologia* **37**: 84-91
- Neta-Sharir I, Isaacson T, Lurie S, Weiss D** (2005) Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. *Plant Cell* **17**: 829-1838
- Neumann P, Koblikova A, Navratilova A, Macas J** (2006) Significant expansion of *Vicia pannonica* genome size mediated by amplification of a single type of giant retroelement. *Genetics* **173**: 1047-1056
- Nishida I, Murata N** (1996) Chilling sensitivity in plants and cyanobacteria: the crucial contribution of membrane lipids. *Annu Rev Plant Physiol Plant Mol Biol* **47**:541-568
- Niyogi KK, Truong TB** (2013) Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr Opin Plant Biol* **16**: 1-8
- Nguyen-Deroche TLN** (2008) Thèse de doctorat : Effets du cuivre et du zinc sur la photosynthèse et le métabolisme d'une cyanobactérie et de diatomées marines du littoral vietnamien; comparaison avec des espèces de la côte atlantique française. Université du Maine, 164 pages

- Nguyen-Deroche TLN, Le TT, Bui TV, Rincé Y, Tremblin G, Morant-Manceau A** (2009) Effects of copper on growth and photosynthesis in marine diatoms : a comparison between species from two different geographical areas. *Cryptogamie. Algologie* **30**: 13
- Nguyen-Deroche TLN, Caruso A, Le TT, Bui TV, Schoefs B, Tremblin G, Morant-Manceau A** (2012) Zinc affects differently growth, photosynthesis, antioxidant enzyme activities and phytochelatin synthase expression of four marine diatoms. *The Scientific World J* doi :10.1100/2012/982957, 15p
- Nesmelova IV, Hackett PB** (2010) DDE transposases: Structural similarity and diversity. *Advanced drug delivery reviews* **62**: 1187-1195
- Nozaki H, Takano H, Misumi O, Terasawa K, Matsuzaki M, Maruyama S, Nishida K, Yagisawa F, Yoshida Y, Fujiwara T, Takio S, Tamura K, Chung SJ, Nakamura S, Kuroiwa H, Tanaka K, Sato N, Kuroiwa T** (2007) A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae*. *BMC Biol* **5**: 28.
- Nymark M, Valle KC, Brembu T, Hancke K, Winge P, Andresen K, Johnsen G, Bones AM** (2009) An integrated analysis of molecular acclimation to high light in the marine diatom *Phaeodactylum tricornutum*. *PLoS ONE* **4**: e7743

## O

- Okamoto H, Hirochika H** (2001) Silencing of transposable elements in plants. *Trends Plant Sci* **6**: 527-534
- Okamoto N, Inouye I** (2005) The katablepharids are a distant sister group of the Cryptophyta: A proposal for Katablepharidophyta divisio nova/ Kathablepharida phylum novum based on SSU rDNA and beta-tubulin phylogeny. *Protist* **156**: 163-79.
- Oliver MJ, Schofield O, Bidle K** (2010) Density dependent expression of a diatom retrotransposon. *Mar Genomics* **3**: 145-150
- O'Malley MA, Simpson AGB, Roger AJ** (2013) The other eukaryotes in light of evolutionary protistology. *Biol Philos* **28**: 299-330
- Oxborough K** (2004). Imaging of chlorophyll *a* fluorescence: theoretical and practical aspects of an emerging technique for the monitoring of photosynthetic performance. *J Exp Bot* **55**: 1195-1205



## P

- Pace JK, Gilbert C, Clark MS, Feschotte C** (2008). Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci USA* **105**: 17023-17028
- Palomeque T, Antonio Carrillo J, Muñoz-López M, Lorite P** (2006) Detection of a mariner-like element and a miniature inverted-repeat transposable element (MITE) associated with the heterochromatin from ants of the genus *Messor* and their possible involvement for satellite DNA evolution. *Gene* **371**: 194-205
- Panchuk II, Volkov RA, Schöffl F** (2002) Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in *Arabidopsis*. *Plant Physiol* **129**: 838-853
- Parfrey LW, Barbero E, Lasser E, Dunthorn M, Bhattacharya D, Patterson DJ, Katz LA** (2006) Evaluating support for the current classification of eukaryotic diversity. *PLoS Genet* **2**: e220
- Park H, Song B, Morel FMM** (2007) Diversity of the cadmium-containing carbonic anhydrase in marine diatoms and natural waters. *Environ Microbiol* **9**: 403-413
- Parker MS, Armbrust EV** (2005) Synergistic effects of light, temperature, and nitrogen source on transcription of genes for carbon and nitrogen metabolism in the centric diatom *Thalassiosira pseudonana* (Bacillariophyceae). *J Phycol* **41**: 1142-1153
- Payne JL, van de Schootbrugge B** (2007) Life in Triassic Oceans: Links between planktonic and benthic recovery and radiation. In *Evolution of Primary Producers in the Sea* P. Falkowski and A.H. Knoll *eds.* Academic Press. Amsterdam. pp. 165-189
- Petrov DA, Schutzman JL, Hartl DL, Lozovskaya ER** (1995) Diverse transposable elements are mobilized in hybrid dysgenesis in *Drosophila virilis*. *Proc Natl Acad Sci USA* **92**: 8050-8054
- Pfaffl MW, Tichopad A, Progomet C, Neuvians TP** (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnol Lett* **26**: 509-515
- Picot S, Wallau GL, Loreto ELS, Heredia FO, Hua-Van A, Capy P** (2008) The *mariner* transposable element in natural populations of *Drosophila simulans*. *Heredity* **101**: 53-59.
- Piegu B, Guyot R, Picault N, Roulin A, Sanyal A, Kim H, Collura K, Brar DS, Jackson S, Wing RA, Panaud O** (2006) Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Res* **16**: 1262-1269

- Pietrokovski S, Henikoff S** (1997) A helix-turn-helix DNA-binding motif predicted for transposases of DNA transposons. *Mol Gen Genet* **254**: 689-695
- Pinto E, Sigaud-Kutner TCS, Leitaó MAS, Okamoto OK, Morse D, Colepicolo P** (2003) Heavy metal-induced oxidative stress in algae. *J Phycol* **39**: 1008-1018
- Plasterk RH, Izsvák Z, Ivics Z** (1999) Resident aliens : the *Tc1/mariner* superfamily of transposable elements. *Trends Genet* **15**: 326-332
- Potvin G, Zhang Z** (2010) Strategies for high-level recombinant protein expression in transgenic microalgae: a review. *Biotechnol Adv* **28**: 910-918
- Price NM, Morel FMM** (1990) Cadmium and cobalt substitution for zinc in a marine diatom. *Nature* **344**: 658-660
- Pujolar JM, Astolfi L, Boscari E, Vidotto M, Barbisan F, Bruson A, Congiu L** (2013) Tana1, a new putatively active Tc1-like transposable element in the genome of sturgeons. *Mol Phylogenet Evol* **66**: 223-232

## R

- Rai LC, Gaur JP** (2001) Algal adaptation to environmental stresses. Physiological, biochemical and molecular mechanisms. Verlag Springer, New York
- Rajadurai M, Poornima EH, Narasimhan SV, Rao NR, Venugopalan VP** (2005) Phytoplankton growth under temperature stress: Laboratory studies using two diatoms from a tropical coastal power station site. *J Ther Biol* **30**: 299-305
- Raven JA, Waite AM** (2004) The evolution of silification in diatoms : inescapable sinking and sinking as escape ? *New Phytol* **162**: 45-61
- Reyes-Prieto A, Weber AP, Bhattacharya D** (2007) The origin and establishment of the plastid in algae and plants. *Annu Rev Genet* **41**: 147-68
- Rech M, Mouget JL, Tremblin G** (2003) Modification of the Hansatech FMS fluorometer to facilitate measurements with microalgal cultures. *Aquatic Botany* **77**: 71-80
- Rech M, Morant-Manceau A, Tremblin G** (2008) Carbon fixation and carbonic anhydrase activity in *Haslea ostrearia* (Bacillariophyceae) in relation to growth irradiance. *Photosynthetica*, **46**: 56-62
- Rech M, Mouget JL, Morant-Manceau A, Rosa P, Tremblin G** (2005) Long-term acclimation to UV radiations: effect on growth, photosynthesis and carbonic anhydrase activity in marine diatoms. *Bot Mar* **48**: 407-420

- Reinfelder JR, Kraepiel AML, Morel FMM** (2000) Unicellular C4 photosynthesis in marine diatom. *Nature* **407**: 996-999
- Richardson JM, Colloms SD, Finnegan DJ, Walkinshaw MD** (2009) Molecular architecture of the *MosI* paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell* **138**: 1096-1108
- Richardson JM, Dawson A, O'Hagan N, Taylor P, Finnegan DJ, Walkinshaw MD** (2006) Mechanism of *MosI* transposition: insights from structural analysis. *EMBO J* **25**: 1324-1334
- Richardson JM, Zhang L, Marcos S, Finnegan DJ, Harding MM, Taylor P, Walkinshaw MD** (2004) Expression, purification and preliminary crystallographic studies of a single-point mutant of *MosI* mariner transposase. *Acta Crystallogr D Biol Crystallogr* **60**: 962-964
- Rigal M, Mathieu OA** (2011) "mille-feuille" of silencing: epigenetic control of transposable elements. *Biochim Biophys Acta* **1809**: 452-458
- Rijstenbil JW** (2003) Effects of UVB radiation and salt stress on growth, pigments and antioxidative defence of the marine diatom *Cylindrotheca closterium*. *Mar Ecol Progress Ser* **254**: 37-48
- Rijstenbil JW, Derksen JWM, Gerringa LJA, Poortvliet TCW, Sandee A, van den Berg M, van Drie J, Wijnholds A** (1994) Oxidative stress induced by copper: defense and damage in the marine planktonic diatom *Ditylum brightwellii*, grown in continuous cultures with high and low zinc levels. *Mar Biol* **119**: 583-590
- Robertson HM** (2002) Evolution of DNA transposons in eukaryotes. *In : Mobile DNA II*, N. L. Craig, R. Craigie, M. Gellert, et A. M. Lambowitz (Eds). American Society of Microbiology Press, Washington, pp. 1093-1110.
- Robertson HM** (1993) The *mariner* transposable element is widespread in insects. *Nature* **362**: 241-245
- Robertson HM** (1997) Multiple *Mariner* transposons in flatworms and hydras are related to those of insects. *J Hered* **88**: 195-201
- Robertson HM, Asplund ML** (1996) *BmmarI*: a basal lineage of the mariner family of transposable elements in the silkworm moth, *Bombyx mori*. *Insect Biochem Mol Biol* **26**: 945-954
- Robertson HM, Lampe DJ** (1995) Recent horizontal transfer of a *mariner* transposable element among and between *Diptera* and *Neuroptera*. *Mol Biol Evol* **12**: 850-862

- Robertson HM, Martos R** (1997) Molecular evolution of the second ancient human mariner transposon, *Hsmar2*, illustrates patterns of neutral evolution in the human genome lineage. *Gene* **205**: 219-228
- Robertson HM, Zumpano KL** (1997) Molecular evolution of an ancient mariner transposon, *Hsmar1*, in the human genome. *Gene* **205**: 203-217
- Robertson HM, Lampe DJ, MacLeod EG** (1992) A mariner transposable element from a lacewing. *Nucleic Acids Res* **20**: 6409
- Roberts K, Granum E, Leegood RC, Raven JA** (2007) C3 and C4 pathways of photosynthetic carbon assimilation in marine diatoms are under genetic, not environmental, control. *Plant Physiol* **145**: 230-235
- Robertson HM, Zumpano KL, Lohe AR, Hartl DL** (1996) Reconstructing the ancient mariners of humans. *Nat Genet* **12**: 360-361
- Rohacek K, Soukupova J, Bartak M** (2008) Chlorophyll fluorescence: a wonderful tool to study plant physiology and plant stress. In *Plant Cell Compartments*. Ed. B. Schoefs, Keral, India
- Ronquist F, Huelsenbeck JP** (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574
- Rosenzweig B, Liao LW, Hirsh D** (1983) Sequence of the *C. elegans* transposable element Tc1. *Nucleic Acids Res* **11**: 4201-4209
- Rouault JD, Casse N, Hermann D, Abdallah Z, Bettaibi A, Kharrat I** (2012) An overview of the Tc1-mariner-IS630 SuperFamily. International Congress of Transposable Elements, ICTE, Saint-Malo, France, April 21-24.
- Roulin A, Piegu B, Wing R, Panaud O** (2008) Evidence of multiple horizontal transfers of the long terminal repeat retrotransposon *RIRE1* within the genus *Oryza*. *The plant J* **53**: 950-959
- Roulin A, Piegu B, Fortune P, Sabot F, D'Hont A, Manicacci D, Panaud O** (2009) Whole genome surveys of rice, maize and sorghum reveal multiple horizontal transfers of the LTR-retrotransposon Route66 in Poaceae. *BMC Evol Biol* **9**: 58
- Round FE, Crawford RM, Mann DG** (1990) *The Diatoms. Biology and morphology of the genera*. Cambridge University Press, Cambridge 747 p
- Rousch JM, Scott EB, Milton RS** (2004) Protein expression during heat stress in thermo-intolerant and thermo tolerant diatoms. *J Exp Mar Biol Ecol* **306**: 231-243

- Rousseau D, Helies-Toussaint C, Moreau D, Raederstorff D, Grynberg A** (2003) Dietary n-3 PUFAs affect the blood pressure rise and cardiac impairments in a hyperinsulinemia rat model in vivo. *Am J Heart and Circulatory Physiol* **285**: 1294-1302
- Rubin E, Levy AA** (1997) Abortive gap repair: underlying mechanism for *Ds* element formation. *Mol Cell Biol* **17**: 6294-6302
- Rubin GM, Spradling AC** (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348-353
- Rubin PM, Loreto EL, Carareto CM, Valente VL** (2011) The copia retrotransposon and horizontal transfer in *Drosophila willistoni*. *Genet Res (Camb)* **31**: 1-6

## S

- Saakov V** (2002) High-temperature stress-related changes in the harmonics  $F_0$ ,  $F_m$  and  $F_v$  of pulse-amplitude modulated fluorescence signals: Locating thermal damage in reaction center in photosystem II. *Doklady Biochemistry and Biophysics* **382**: 4-9
- Sambrook J, Fritsch EF, Maniatis T** (1989) *Molecular Cloning: A laboratory Manual*, Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY, 1989
- Sasakura Y, Awazu S, Chiba S, Kano S, Satoh N** (2003) Application of *Minos*, one of the *Tc1 / mariner* superfamily transposable elements, to ascidian embryos as a tool for insertional mutagenesis. *Gene* **308**: 11-20
- Scarpeci TE, Zanol MI, Valle EM** (2008) Investigating the role of plant heat shock proteins during oxidative stress. *Plant Signal Behav* **3**: 856-857
- Schaack S, Gilbert C, Feschotte C** (2010) Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* **25**: 537-546
- Schnerr H, Niessen L, Vogel RF** (2001) Real time detection of the *tri5* gene in *Fusarium* species by LightCycler-PCR using SYBR Green I for continuous fluorescence monitoring. *Int J Food Microbiol* **71**: 53-61
- Schrader SM, Wise RR, Wacholtz WF, Ort DR, Sharkey TD** (2004) Thylakoid membrane responses to moderately high leaf temperature in Pima cotton. *Plant, Cell and Environ* **27**: 725-735

- Schubert H, Tremblin G, Robert JM, Sagert S, Rincé Y** (1995) In-vivo measurement of photosynthesis of *Haslea ostrearia* Simonsen in relation to marennine content. *Diatom Res* **10**: 341-349
- Schulz-Raffelt M, Lodha M, Schroda M** (2007) Heat shock factor 1 is a key regulator of the stress response in *Chlamydomonas*. *Plant J* **52**: 286-295
- Shao H, Tu Z** (2001) Expanding the diversity of the *IS630-Tc1-mariner* superfamily : discovery of a unique DD37E transposon and reclassification of the DD37D and DD39D transposons. *Genetics* **159**: 1103-1115
- Sharkey T** (2005) Effects of moderate heat stress on photosynthesis: importance of thylacoid reactions, Rubisco deactivation, reactive oxygen species, and thermotolerance provided by isoprene. *Plant, Cell and Environment* **28**: 269-277
- Shepherd CT, Lauter ANM, Scott P** (2009) Determination of Transgene Copy Number by Real-Time Quantitative PCR. *Methods Mol Biol* **526**: 129-134
- Shi Y, Mosser DD, Morimoto RI** (1998) Molecular chaperones as HSF1-specific transcriptional repressors. *Genes Dev* **12**: 654-666
- Siaut M, Heijde M, Mangogna M, Montsant A, Coesel S, Allen A, Manfredonia A, Falciatore A, Bowler C** (2007) Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*. *Gene* **406**: 23-35
- Silva JC, Kidwell MG** (2000) Horizontal transfer and selection in the evolution of *P* elements. *Mol Biol Evol* **17**: 1542-1557
- Silva JC, Kidwell MG** (2004) Evolution of *P* elements in natural populations of *Drosophila willistoni* and *D. sturtevantii*. *Genetics* **168**: 1323-1335
- Silva JC, Bastida F, Bidwell SL, Johnson PJ, Carlton JM** (2005) A potentially functional mariner transposable element in the protist *Trichomonas vaginalis*. *Mol Biol Evol* **22**: 126-134
- Simon N, Cras AL, Foulon E, Lemée R** (2009) Diversity and evolution of marine phytoplankton. *C R Acad Sci Biologies* **332**: 159-170
- Sinzelle L, Izsvak Z, Ivics Z** (2009) Molecular domestication of transposable elements: from detrimental parasites to useful host genes. *Cell Mol Life Sci* **66**: 1073-1093
- Sinzelle L, Jégot G, Brillet B, Rouleux-Bonnin F, Bigot Y, Augé-Gouillou C** (2008) Factors acting on *Mos1* transposition efficiency. *BMC Mol Biol* **9**: 106

- Sinzelle L, Pollet N, Bigot Y, Mazabraud A** (2005) Characterization of multiple lineages of Tc1-like elements within the genome of the amphibian *Xenopus tropicalis*. *Gene* **349**: 187-196
- Slotkin RK, Martienssen R** (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics* **8**: 272-285
- Sobecky PA, Hazen TH** (2009) Horizontal gene transfer and mobile genetic elements in marine systems. *Methods in Molecular Biology* **532**: 435-453
- Smetacek V** (1999) Diatoms and the ocean carbon cycle. *Protist* **150**: 25-32
- Sorhannus U** (2011) Evolution of antifreeze protein genes in the diatom genus *fragilariopsis*: evidence for horizontal gene transfer, gene duplication and episodic diversifying selection. *Evol Bioinform Online* **7**: 279-289
- Speziale B J, Schreiner S P, Giammatteo PA, Scxhindler J E** (1984). Comparison of N,N-dimethylformamide, dimethyl sulfoxide, and acetone for extraction of phytoplankton chlorophyll. *Canadian Journal of Fisheries and Aquatic Sciences* **41**: 1519-1522
- Staubert JL, Florence TM** (1990) Mechanism of toxicity of zinc to the marine diatom *Nitzschia closterium*. *Mar Biol* **105**: 519-524
- Stengel DB, Connan S, Popper ZA** (2011) Algal chemodiversity and bioactivity: sources of natural variability and implications for commercial application. *Biotech Adv* **29**: 483-501
- Stockinger EJ, Gilmour SJ, Thomashow MF** (1997) *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci USA* **94**: 1035-1040
- Strzepeck RJ, Harrison PJ** (2004) Photosynthetic architecture differs in coastal and oceanic diatoms. *Nature* **431**: 689-692
- Sumitani M, Lee JM, Hatakeyama M, Oishi K** (2002) Cloning and characterization of Acmar1, a mariner-like element in the asiatic honey bee, *Apis cerana japonica* (Hymenoptera, Apocrita). *Arch Insect Biochem Physiol* **50**: 183-190
- Sunda WG, Huntsman S** (1997) Interrelated influence of iron, light and cell size on marine phytoplankton growth. *Nature* **390**: 389-392
- Sundaresan V** (1996) Horizontal spread of transposons mutagenesis: new uses for old elements. *Trends Plant Sci* **1**: 184-190
- Sung DY, Vierling E, Guy CL** (2001) Comprehensive expression profile analysis of the *Arabidopsis* Hsp70 gene family. *Plant Physiol* **126**: 789-800

## T

- Tachibana M, Allen AE, Kikutani S, Endo Y, Bowler C, Matsuda Y** (2011) Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth Res* **109**: 205-221
- Takahashi R, Morita Y, Nakayama M, Kanazawa A, Abe J** (2012) An Active CACTA-Family transposable element is responsible for flower variegation in wild soybean *Glycine soja*. *The Plant Genome* **5**: 62-70
- Tamura K, Dudley J, Nei M, Kumar S** (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Bio Evol* **24**:1596-1599
- Tang SX, Hass CG, Knapp SJ** (2006) *Ty3/gypsy*-like retrotransposon knockout of a 2-methyl-6-phytyl-1,4-benzoquinone methyltransferase is non-lethal, uncovers a cryptic paralogous mutation, and produces novel tocopherol (vitamin E) profiles in sunflower. *Theor Appl Genet* **113**: 783-799
- Tarchini R, Biddle P, Wineland R, Tingey S, Rafalski A** (2000) The complete sequence of 340 kb of DNA around the rice *Adh1-adh2* region reveals interrupted colinearity with maize chromosome 4. *Plant Cell* **12**: 381-391
- Tenaillon MI, Hollister JD, Gaut BS** (2010) A triptych of the evolution of plant transposable elements. *Trends Plant Sci* **15**: 471-478
- Thiriet-Rupert S** (2013) Rapport de Master II : Réponse de la microalgue *Amphora acutiuscula* à des stress thermiques et métalliques. 42 pages
- Tien CJ, Sigee DC, White KN** (2005) Copper adsorption kinetics of cultured algal cells and freshwater phytoplankton with emphasis on cell surface characteristics. *J Appl Phycol* **17**: 379-389
- Tirichine L, Bowler C** (2011) Decoding algal genomes: tracing back the history of photosynthetic life on Earth. *Plant J* **66**: 45-57
- Tomita S, Sohn BH, Tamura T** (1997) Cloning and characterization of a mariner-like element in the silkworm, *Bombyx mori*. *Genes & Genetic Systems* **72**:219-228
- Tortell PD** (2000) Evolutionary and ecological perspectives on carbon acquisition in phytoplankton. *Limnol Oceanogr* **45**: 744-750
- Torti C, Gomulski LM, Moralli D, Raimondi E, Robertson HM, Capy P, Gasperi G, Malacrida AR** (2000) Evolution of different subfamilies of *mariner* elements within the medfly genome inferred from abundance and chromosomal distribution. *Chromosoma* **108**: 523-532



**Tosi LRO, Beverley SM** (2000) *Cis* and *trans* factors affecting *Mos1* mariner evolution and transposition *in vitro*, and its potential for functional genomics. *Nucleic Acids Res* **28**: 784-790

**Tsvetkova NM, Horváth I, Török Z, Wolkers WF, Balogi Z, Shigapova N, Crowe LM, Tablin F, Vierling E, Crowe JH, Vigh L** (2002) Small heat-shock proteins regulate membrane lipid polymorphism. *Proc Natl Acad Sci USA* **99**: 13504-13509

**Turcotte K, Srinivasan S, Bureau T** (2001) Survey of transposable elements from rice genomic sequences. *Plant J* **25**: 169-179

**Turmel M, Gagnon MC, O'Kelly CJ, Otis C, Lemieux C** (2009) The chloroplast genomes of the green algae *Pyramimonas*, *Monomastix*, and *Pycnococcus* shed new light on the evolutionary history of prasinophytes and the origin of the secondary chloroplasts of euglenids. *Mol Biol Evol* **26**: 631-648

## U

**Uemura M, Tominaga Y, Nakagawara, Shigematsu, Minami A, Kawamura Y** (2006) Responses of the plasma membrane to low temperatures. *Physiol Plant* **126**: 81-89

## V

**Van Berkel J, Salamini F, Gebhardt C** (1994) Transcripts accumulating during cold storage of potato (*Solanum tuberosum* L.) tubers are sequences related to stress-responsive genes. *Plant Physiol* **104**: 445-452

**Van Buskirk HA, Thomashow MF** (2006) Arabidopsis transcription factors regulating cold acclimation. *Physiol Plant* **126**: 72-80

**Vandesompele J, Kubista M, Pfaffl W** (2009) Reference gene validation software for improved normalization. In "Real-time PCR: Current Technology and Applications", Logan J, Edwards K, Saunders N (eds), Caister Academic Press, Norfolk, UK, ISBN 978-1-904455-39-4 pp 47-64

**Vardi A, Thamatrakoln K, Bidle KD, Falkowski PG** (2008) Diatom genomes come of age. *Genome Biol* **9**: 245

**Vardi A, Van Mooy BA, Fredricks HF, Popenoerf KJ, Ossolinski JE, Haramaty L, Bidle KD** (2009) Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* **326**: 861-865

- Vaucheret H, Fagard M** (2001) Transcriptional gene silencing in plants: targets, inducers and regulators. *Trends Genet* **17**: 29-35
- Vavilin DV, Ducruet JM, Matorin DN, Venediktov PS, Rubin AB** (1998). Membrane lipid peroxidation, cell viability and Photosystem II activity in the green alga *Chlorella pyrenoidosa* subjected to various stress conditions. *J Photochem Photobiol Biol* **42**: 233-239
- Vieira C, Aubry P, Lepetit D, Biémont C** (1998) A temperaturecline in copy number for 412 but not roo/B104 retrotransposons in populations of *Drosophila simulans*. *Proc Biol Sci* **265**: 1161-1165
- Vierling E** (1991) The roles of heat shock proteins in plants. *Ann Rev Plant Physiol Plant Mol Biol* **42**: 579-620
- Vitte C, Panaud O** (2005) LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet. Genome Res* **110**: 91-107
- Vos JC, Plasterk RH** (1994) *Tc1* transposase of *Caenorhabditis elegans* is an endonuclease with a bipartite DNA binding domain. *EMBO J* **13**: 6125-6132

## W

- Wallau GL, Hua-Van A, Capy P, Loreto EL** (2011) The evolutionary history of *mariner*-like elements in *Neotropical drosophilids*. *Genetica* **139**: 327-338
- Wallau GL, Ortiz MF, Loreto EL** (2012) Horizontal transposon transfer in eukarya: detection, bias, and perspectives. *Genome Biol Evol* **4**: 689-699
- Wang J, Staten R, Miller TA, Park Y** (2005) Inactivated *mariner*-like elements (*MLE*) in pink bollworm, *Pectinophora gossypiella*. *Insect Molecular Biology* **14**: 547-553
- Wang SC, Schnell RA, Lefebvre PA** (1998) Isolation and characterization of a new transposable element in *Chlamydomonas reinhardtii*. *Plant Mol Biol* **38**: 681-687
- Waters ER, Rioflorido I** (2007) Evolutionary analysis of the small heat shock proteins in five complete algal genomes. *J Mol Evol* **65**: 162-174
- Watson D, Daume S, Prince J, Beazley L, Knott B** (2004) The influence of light intensity on the density of different diatoms as feed for juvenile greenlip abalone (*Haliotis laevis*). *Aquaculture* **235**: 345-359
- Wagh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BB, Powel W** (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome

revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol Gen Genet* **253**: 687-694

**Whelan JA, Russell NB, Whelan MA** (2003) A method for the absolute quantification of cDNA using real-time PCR. *J Immunol Methods* **278**: 261-269

**Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH** (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* **8**: 973-982

**Wilhelm C, Büchel C, Fisahn J, Goss R, Jakob T, La Roche J, Lavaud J, Lohr M, Riebesell U, Stehfest K, Valentin K, Kroth PG** (2006) The regulation of carbon and nutrient assimilation in diatoms is significantly different from algae. *Protist* **157**: 91-124

**Wolfe-Simon FD, Grzebyk S, Schofield S, Falkowski PG** (2005) The role and evolution of superoxide dismutase in algae. *J Phycol* **41**: 453-465

**Wong SW, Leung PT, Djurisić AB, Leung KM** (2010) Toxicities of nano zinc oxide to five marine organisms: influences of aggregate size and ion solubility. *Anal Bioanal Chem* **396**: 609-618

**Woodcock DM, Lawler CB, Linsenmeyer ME, Doherty JP, Warren WD** (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J Biol Chem* **272**: 7810-7816

**Wright S, Finnegan D** (2001) Genome evolution: sex and the transposable element. *Cur Biol* **11**: 296-298

**Wu TM, Hsu YT, Lee TM** (2009) Effects of cadmium on the regulation of antioxidant enzyme activity, gene expression, and antioxidant defenses in the marine macroalga *Ulva fasciata*. *Botanical Studies* **50**: 25-34

**Wu-Scharf D, Jeong B, Zhang C, Cerutti H** (2000) Transgene and transposon silencing in *Chlamydomonas reinhardtii* by a DEAH-box RNA helicase. *Science* **290**: 1159-1162

## Y

**Yang G, Hall TC** (2003) MDM-1 and MDM-2: two *Mutator*-derived MITE families in rice. *J Mol Evol* **56**: 255-264

- Yang G, Nagel DH, Feschotte C, Hancock CN, Wessler SR** (2009) Tuned for transposition: molecular determinants underlying the hyperactivity of a *Stowaway* MITE. *Science* **325**: 1391-1394
- Yang G, Weil CF, Wessler SR** (2006) A rice Tc1/mariner-like element transposes in yeast. *Plant Cell* **18**: 2469-2478
- Yongmanitchai W, Ward OP** (1991) Growth of and omega-3 fatty acid production by *Phaeodactylum tricornutum* under different culture conditions. *App Environ Microbiol* **57**: 419-425
- Yu GX, Wise RP** (2000) An anchored AFLP- and retrotransposon-based map of diploid *Avena*. *Genome* **43**: 736-749
- Yuan JS, Wang D, Stewart CN Jr** (2008) Statistical methods for efficiency adjusted real-time PCR quantification. *Biotechnol J* **3**: 112-123
- Yun MS, Lee SH, Chung IK** (2010) Photosynthetic activity of benthic diatoms in response to different temperatures. *J Appl Phycol* **22**: 559-562

## **Z**

- Zakharkin SO, Willis RL, Litvinova OV, Jinwal UK, Headley VV, Benes H** (2004) Identification of two mariner-like elements in the genome of the mosquito *Ochlerotatus atropalpus*. *Insect Biochem Mol Biol* **34**: 377-386
- Zemach A, McDaniel IE, Silva P, Zilberman D** (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**: 916-919
- Zhang L, Dawson A, Finnegan DJ** (2001) DNA-binding activity and subunit interaction of the *mariner* transposase. *Nucleic Acids Res* **29**: 3566-3575
- Zhang X, Jiang N, Feschotte C, Wessler SR** (2004) *PIF*- and *Pong*-like transposable elements: distribution, evolution and relationship with *Tourist*-like miniature inverted-repeat transposable elements. *Genetics* **166**: 971-986
- Zhou MB, Lu JJ, Zhong H, Tang KX, Tang DQ** (2010) Distribution and polymorphism of *mariner*-like elements in the *Bambusoideae* subfamily. *Plant Syst Evol* **289**: 1-11
- Zhou MB, Zhong H, Tang DQ** (2011) Isolation and characterization of seventy-nine full-length mariner-like transposase genes in the *Bambusoideae* subfamily. *J Plant Res* **124**: 607-617

## ANNEXES

### Annexe 1 : Eau de mer artificielle (EMA)

L'eau de mer artificielle est préparée d'après le protocole de Harrison *et al.* (1980) avec quelques modifications. Les solutions I et II sont préparées séparément puis la solution I est complétée par l'ajout de vitamines filtrées sur filtre 0,22 µm et les métaux et nutriments sont ajoutés à la solution II avant autoclavage. Ensuite ces deux solutions sont autoclavées séparément. Après 24 heures, elles sont mélangées sous une hotte à flux laminaire puis l'EMA est conservée à 4 °C. Pour éviter qu'*Amphora acutiuscula* n'adhère aux parois des fioles Erlenmeyer, la teneur en Ca de l'EMA a été réduite de 9,4 mM à 0,25mM. En parallèle, la concentration en NaCl a été augmentée de 363 à 368 mM.

		Concentration finale (M)
Solution I : sels anhydres	NaCl	$3,63.10^{-1}$
	Na <sub>2</sub> SO <sub>4</sub>	$2,50.10^{-2}$
	KCl	$8,04.10^{-3}$
	NaHCO <sub>3</sub>	$2,07.10^{-3}$
	KBr	$7,25.10^{-4}$
	H <sub>3</sub> BO <sub>3</sub>	$3,72.10^{-4}$
	NaF	$6,57.10^{-5}$
Solution II : sels hydratés	MgCl <sub>2</sub> , 6H <sub>2</sub> O	$4,12.10^{-2}$
	CaCl <sub>2</sub> , 2H <sub>2</sub> O	$9,40.10^{-3}$
	SrCl <sub>2</sub> , 6H <sub>2</sub> O	$8,20.10^{-5}$
Métaux	FeCl <sub>3</sub> , 6H <sub>2</sub> O	$6,56.10^{-6}$
	Na <sub>2</sub> EDTA, 2H <sub>2</sub> O	$1,49.10^{-5}$
	CoSO <sub>4</sub> , 7H <sub>2</sub> O	$5,69.10^{-8}$
	Na <sub>2</sub> MoO <sub>4</sub> , 7H <sub>2</sub> O	$6,10.10^{-9}$
	Na <sub>2</sub> SeO <sub>3</sub>	$1,00.10^{-9}$
	NiCl <sub>2</sub> , 6H <sub>2</sub> O	$6,30.10^{-9}$
	ZnSO <sub>4</sub> , 7H <sub>2</sub> O	$2,54.10^{-7}$
	CuSO <sub>4</sub> , 5H <sub>2</sub> O	$3,92.10^{-8}$
	MnSO <sub>4</sub> , H <sub>2</sub> O	$2,42.10^{-6}$
Nutriments majeurs	FeNH <sub>4</sub> -citrate	$1,37.10^{-6}$
	NaNO <sub>3</sub>	$5,49.10^{-4}$
	NaH <sub>2</sub> PO <sub>4</sub> , 2H <sub>2</sub> O	$2,10.10^{-5}$
	Na <sub>2</sub> SiO <sub>3</sub> .9H <sub>2</sub> O	$1,05.10^{-4}$
Vitamines	thiamine	$2,97.10^{-7}$
	biotine	$4,09.10^{-9}$
	B12	$1,47.10^{-9}$
	acide nicotinique	$6,50.10^{-10}$
	acide folique	$9,06.10^{-9}$
	thymine	$9,52.10^{-7}$
	ca-d-pantothenate	$8,39.10^{-9}$
	inositol	$1,11.10^{-6}$

## Annexe 2 : Axénie des cultures d'*Amphora acutiuscula*

Afin d'éliminer tout risque de contamination des cultures d'*Amphora acutiuscula* par des bactéries et des champignons, une culture dense d'*A. acutiuscula* a été traitée par un mélange d'antibiotiques et d'antifongique (10.000 unités pénicilline G, 10 mg streptomycine, 25 µg amphotérycine B, par mL; Sigma-Aldrich A5955, Saint Quentin Fallavier, France) à raison de 1,25 mL pour 125 mL de culture. Pour un retour à la phase exponentielle de croissance, la culture est placée pendant 5 jours dans des conditions de culture standard. Dans certains cas, l'obtention de l'axénie peut nécessiter l'ajout de plusieurs doses d'antibiotiques/ antifongique, en suivant la même procédure.

L'éventuelle présence de bactéries et de champignons microscopiques a été contrôlée par un test réalisé par inoculation d'un mL de culture de microalgues dans les milieux FAG et FG (composition ci-dessous) gélosés. Le premier milieu FAG est favorable à la croissance des mycètes et le deuxième (FG) est favorable au développement bactérien. Les boîtes sont placées à l'obscurité durant 5 à 6 jours à température ambiante pour inhiber la croissance des microalgues.

L'absence de colonies bactériennes et fongiques sur les boîtes 5 jours après l'inoculation permet de conclure à l'absence de bactérie et de champignon dans la culture de microalgues. Une comparaison avec deux autres boîtes de Pétri, l'une ensemencée par une culture non-axénique et l'autre par de l'eau de mer, permet de confirmer que la culture mère est bien axénique.

A chaque manipulation, le col des fioles a été passé dans la flamme d'une lampe à alcool afin de contribuer à maintenir l'axénie. Dans le cas où une contamination est observée au cours d'une culture, le protocole ci-dessus est à nouveau appliqué.

Composition	Milieu FAG	Milieu FG
Bacto-peptone	3 g	4 g
Extrait de levure	1 g	0,5 g
NH <sub>4</sub> SO <sub>4</sub>	1 g	0 g
Glycérophosphate Na	25 mg	25 mg
EDTA ferrique	6 mg	6 mg
Eau de mer artificielle	50 % qsp 1 L	75 % qsp 1L
pH	7,4 -7,8	7,4 -7,8











	2100	2120	2140	2160
Aacutmar1.4	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2130	
Aacutmar1.12	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2130	
Aacutmar1.3	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2114	
Aacutmar1.7	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2111	
Aacutmar1.2	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2114	
Aacutmar1.1	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2114	
Aacutmar1.9	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2114	
Aacutmar1.6	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2116	
Aacutmar1.10	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2113	
Aacutmar1.11	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2114	
Aacutmar1.5	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTCGGTCCAGGCGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2132	
Aacutmar1.8	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTCGGTCCAGGCGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2094	
ConsAacutmar1	: TTTGTACCAGTTCCTAGTGGGCTCTTGTTTGCCTGACACGGGA	CTGAGGTAATAATTCGGCCAGAGGGGGG	: 2114	



## Annexe 4 : Alignement de 12 ORF d'*Aacutmar1*.

		20	40	60	80	
Aacutmar1.4	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSANGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.12	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSANGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.3	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.7	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.9	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.2	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.1	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.6	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.10	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	81		
Aacutmar1.11	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	82		
Aacutmar1.5	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		
Aacutmar1.8	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	76		
ConsAacutmar1	:	MRAMPPLFRSAGTLHTTIKMDQENLEPNNAKRVKTKELTDNERIAVVSMLLGMSAKGPLPHGSFTDVAKKFGVPRPCVGELNAKAT	:	84		

		*	100	*	120	*	140	*	160																																																																			
Aacutmar1.4	:	TS	R	I	N	G	K	V	D	B	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168		
Aacutmar1.12	:	TS	R	I	N	G	K	V	D	B	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168		
Aacutmar1.3	:	TS	R	S	N	G	K	V	D	B	A	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168	
Aacutmar1.7	:	TS	R	S	N	G	K	V	D	B	A	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168	
Aacutmar1.9	:	TS	R	S	N	G	K	V	D	B	A	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168	
Aacutmar1.2	:	T	P	R	S	N	G	K	V	D	B	A	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168
Aacutmar1.1	:	TS	R	S	N	G	K	V	D	B	A	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168	
Aacutmar1.6	:	TS	Q	I	N	G	K	V	D	B	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	168		
Aacutmar1.10	:	TS	Q	I	N	G	K	V	D	B	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	165		
Aacutmar1.11	:	TS	Q	I	N	G	K	V	D	B	E	I	S	K	K	H	E	R	G	R	K	M	L	D	I	D	A	M	E	A	V	Q	V	V	S	V	K	S	H	G	L	A	A	I	G	V	P	K	T	T	L	L	Y	R	K	F	L	V	R	H	S	N	A	L	K	P	T	L	P	E	H	K	:	166		
Aacutmar1.5	:	TS	W	I	N	G	K	V	D	B	E																																																																	

		*	180	*	200	*	220	*	240	*				
Aacutmar1.4	:	AARLEFAMSKVD	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.12	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.3	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.7	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.9	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.2	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.1	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.6	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	251
Aacutmar1.10	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	248
Aacutmar1.11	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	249
Aacutmar1.5	:	AARLEFAMSKVD	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252
Aacutmar1.8	:	AARLEFAMSKVD	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	244
ConsAacutmar1	:	AARLEFAMSKID	PNTVRNGS	CFKYL	PMLDE	VHVDEK	WFHMT	KDNQV	FILSKN	KEPPKRTV	QHKCYIGKVMFLCAQ	APRMVDG	:	252

	260	*	280	*	300	*	320	*
Aacutmar1.4	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.12	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.3	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.7	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.9	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.2	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.1	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.6	: MGDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYCCYMMKEK	VIPAIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 335
Aacutmar1.10	: MGDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYCCYMMKEK	VIPAIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 332
Aacutmar1.11	: MGDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYCCYMMKEK	VIPAIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 333
Aacutmar1.5	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336
Aacutmar1.8	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYHGYMLDK	VISAIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 328
ConsAacutmar1	: MWDGKIGIWF	FGYVGEAQRFSK	HRPRGAPVWKNDSV	CEVYRRYMLDKVIP	AIIDKFFPMRYLKG	GVKVIQQDGS	TAHLPGDDEW	: 336

	340	*	360	*	380	*	400	*	420	
Aacutmar1.4	:	NEAMEILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.12	:	NEAMEILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.3	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.7	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.9	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.2	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.1	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						
Aacutmar1.6	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	419						
Aacutmar1.10	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAMEDAYWKYKANKLNRMWLTYSQCLNMII	:	416						
Aacutmar1.11	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	417						
Aacutmar1.5	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	417						
Aacutmar1.8	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	407						
ConsAacutmar1	:	NEAMDILKRESRINIRKLYTQCAQSPDNINDLAFRSIQCLLYEAAPTTFALIKAVEDAYWKYKANKLNRMWLTYSQCLNMII	:	420						

[illegible]

## Annexe 5 : Alignement de *Re-4* et la séquence de sa vérification *BATG/VRe-4*.

```

Re-4      : CGTGTTCACGGCGCTCTGCGAAGCGACCGTGGCGGGCATCGAGCGCGATTTCGAAATCAAATACGGTGTCTGGTGGCAGCGGCTCTTCCGTGCATCA : 95
BATG/VRe4-1 : ----- : -

Re-4      : GCGCAGGAGACAAATTGCCAGTCAAAAGTCGAGCAATGTTGGCGTTGCGGAAACAAGTGGCAGTATCCAAGACGAGTACTCACCGTATCATAGCA : 190
BATG/VRe4-1 : ----- : -

Re-4      : TGCACATTTGGGCCATATACTTTGCTAGAAATTTCTGACCGCCACTCCGAAAACAATATTCTCGTGACGGACATCGTTTTACCTACTCCCTCTGG : 285
BATG/VRe4-1 : ----- : -

Re-4      : TCCAAATTAACATCACTATCGGTCTCTCTTAAATAAAATTTTCAATATCACCCCTTCCTATTAGTAGCTATATCATTGTAATCTACTACAAA : 380
BATG/VRe4-1 : ----- : -

Re-4      : AAGAGCTACATATCCATGTATATATTGAATATTCAATTTGAACAGAAGCAATGAAAAGGGGGCTAAAAGCATGCGCTGCCAATGCCTCTTTT : 475
BATG/VRe4-1 : -----TTGAACAGAAGCAATGAAAAGGGGGCTAAAAGCATGCGCTGCCAATGCCTCTTTT : 57

Re-4      : CGATCAGCTGGAACTCTGCACACAATTGTCAAAATGGACAGGAGAAGCTTGACAACCAAAATACCAAGAGATTAAACAAAGAACTCACCAAC : 570
BATG/VRe4-1 : TGAATCAGGGAACTTGCACACAATTGTCAAAATGGACAGGAGAAGCTTGACAACCAAAATACCAAGAGG----- : 127

Re-4      : ATGAAAGCATTTGAGTGGTTTCCATGCTGTTAGGAATGTCTGCAAAAGGCCCTCTTCCCATGGTAGTTTACAGATGTGGCTAAAAAATTTGA : 665
BATG/VRe4-1 : ATGAAAGCATTTGAGTGGTTTCCATGCTGTTAGGAATGTCTGCAAAAGGCCCTCTTCCCATGGTAGTTTACAGATGTGGCTAAAAAATTTG : 221

Re-4      : GGCACACCCCGCTGTGTAGAGAGCTGTGGGCCAAGGCTACAACCTCTGGTTCAATGGCAAAGTGGATGAAGATGAGTCAAAATCAAGAAAC : 760
BATG/VRe4-1 : ATGTGACCCCTGTGTGTCTGAGAGCTGTGGGCCAAGGCTACAACCTCACATCCAAATGGCAAAGTGGATGAAGATGAGTCAAAATCAAGAAAC : 316

Re-4      : ATGAGAGAGCAAGAAATATGCTGTGGATATCGATGCAATGATGAAGCTGTGCAACAAGTTCCTCGCTCAATGAAGAGTCTTACCATTGGTTG : 855
BATG/VRe4-1 : ATGAGAGAGCAAGAAATATGCTGTGGATATCGATGCAATGATGAAGCTGTGCAACAAGTTCCTCGCTCAATGAAGAGTCTTACCATTGGTTG : 411

Re-4      : GCAGACAGATAGGAGTTTCCAAGACAACCTTTGGCATACAGAAAGGAGCTTCTGTTCAGACACGCAATGCCCTCAAGCCACACTCAGCTCC : 950
BATG/VRe4-1 : GCAGACAGATAGGAGTCCCAAGACAACCTTTGGCATACAGAAAGGAGTCTTCTGTTCAGACATGCAATGCCCTCAAGCCACACTCAGCTCC : 506

Re-4      : TGAACACAAAGCAGCAAGACTTCAATTTGCCATGACCAAGTAGACCCAAATATGGTAAGGAATGGAAGTCCGAAGTACTTACCAATGTAGATG : 1045
BATG/VRe4-1 : TGAACACAAAGCAGCAAGACTTCAATTTGCCATGACCAAGTAGACCCAAATATGGTAAGGAATGGAAGTCCGAAGTACTTACCAATGTAGATG : 601

Re-4      : AGGTTTCATGTTGATGAGAAGTGGTTCCATGACCAAGCACAATCAAGTGTTCATCTTGTCAAGACCGAGCAACCCCAAGAGGACTGTGCAG : 1140
BATG/VRe4-1 : AGGTTTCATGTTGATGAGAAGTGGTTCCATGACCAAGCACAATCAAGTGTTCATCTTGTCAAGACCGAGCAACCCCAAGAGGACTGTGCAG : 696

Re-4      : CACAAGGTTTACATTGGCAAAGTGATGTTCTTGTGTGCACAAGCAAGGCCAAGATGGTAGATGGCGAATGTGGGATGCAAGATTGGCATTG : 1235
BATG/VRe4-1 : CACAAGGTTTACATTGGCAAAGTGATGTTCTTGTGTGCACAAGCAAGGCCAAGATGGT-----GAATGTGGGATGCAAGCTTGGCATTG : 782

Re-4      : GCCATTTGGATATGTTGGACAGACACAAAGATCAAAACACAGACCAGAGGAGGACCGGTTTGGCAAAATGAGAGCGTTAAACAGAGAGGTG : 1330
BATG/VRe4-1 : GCCCTTTGGATATGTTAGGACAGACACAAAGATCAAAACACAGACCAGAGGAGGACCGGTTTGGCAAAATGAGAGCGTTAAACAGAGAGGTG : 877

Re-4      : ACTGCCATATCATGTTTGACAAAGGTCATTCTGCCATTATAGCAAGTTCCCATGACGATCTTGAAGAGTGTGCTGAAATGGGGCTGCAAGATCCAGCAAGAC : 1422
BATG/VRe4-1 : AGAAACATATATGCTTTGACAAAGGTCATTCTGCCATTATAGCAAGTGTGCTGAAATGGGGCTGCAAGATCCAGCAAGAC : 972

Re-4      : AATGCAAAATCCCATATTCCTGACAATGCTGAAGAATGGTTAGATGCCATGCAAAAGCATTACGATGAGTACAAGCTGAATATTCAGATATATAC : 1517
BATG/VRe4-1 : GGTGCAAAATCCCATATTCCTGACAATGCTGAAGAATGGTTAGATGCCATGCAAAAGCATTACGATGAGTACAAGCTGAATATTCAGATATATAC : 1067

Re-4      : TCAACCGGCTCAATCACCAGACACAACATCAATGATTGGCATTTCTTAGAAGCATCCAAATGCTTTATTATGAAGCAGCACCAACAACGACT : 1612
BATG/VRe4-1 : TCAACCGGCTCAATCACCAGACACAACATCAATGATTGGCATTTCTTAGAAGCATCCAAATGCTTTATTATGAAGCAGCACCAACAACGACT : 1162

Re-4      : TTGCTCTCATCAAGAAAGTTGAGGATGCATATTGGAAGTATCCAGTTAACAAGCTCAATCAGATGTTTGAATTTACCAAGAGCTGCTTAACATGA : 1707
BATG/VRe4-1 : TTGCTCTCATCAAGAAAGTTGAGGATGCATATTGGAAGTATCCAGTTAACAAGCTCAATCAGATGTTTGAATTTACCAAGAGCTGCTTAACATGA : 1257

```

Re-4 : 1720 \* 1740 \* 1760 \* 1780 \* 1800 : 1802  
BATG/VRe4-1 : TCATTGAACCGAGGGAGACAACCTTTACTCAATTCTCTCATGAACAAAGAGGGCTTGGGAAACCGGCCGAAGTCTCTCAGAGGAGGAATAC : 1352

Re-4 : \* 1820 \* 1840 \* 1860 \* 1880 \* 1900 : 1897  
BATG/VRe4-1 : TTCATCAAAATGGAACAAAGAACAGAACAGACCAAGATATTGACTATTTTGACATTAGTCTTTCCAATTGTGGTTTCCTTGAATCAGGAGACCT : 1447

Re-4 : \* 1920 \* 1940 \* 1960 \* 1980 \* : 1992  
BATG/VRe4-1 : AAGCTACTCGGCTAGACCATAATTCTCAGGCTACATCAAAACACAGCTAACTGCGACTGGTCCATCTGGGTTCTTTCGATTTTCAGCTGGCTCATT : 1542

Re-4 : 2000 \* 2020 \* 2040 \* 2060 \* 2080 \* : 2087  
BATG/VRe4-1 : TGTCCAACAAAAGGTTTCAGGTTAGGAAAAAAGCAAGCTGACAGCACAACCTTCTGGCTCTTGGACCTCTTCTTATCGCCACACTCTTTCATT : 1637

Re-4 : 2100 \* 2120 \* 2140 \* 2160 \* 2180 : 2182  
BATG/VRe4-1 : CCATTTCGAAAGATAGCCCTTGGTGTGTGTGCCTCTCTTGTGTCTTTTCTCTTCATCCCTTTAATCTTCTTATTCCTTGTATCATGTGGAATT : 1732

Re-4 : \* 2200 \* 2220 \* 2240 \* 2260 \* 2280 : 2277  
BATG/VRe4-1 : CTAATTGGAAGACAGCTTGACGCCGACAGCAGCTTGGGAATATGAAATGCAAGTATCTCGGAAACGGGAGGCTGGTCTCGTCGACGCTGAGGCT : 1827

Re-4 : 2300 \* 2320 \* 2340 \* 2360 : 2372  
BATG/VRe4-1 : CGCTCTCGATTGTACGCACAACCACTTGTCTCCGCCGATCTCAGCGCAGCAGGAAAAACGCCAATTATGGTATGCAGCTTGTATTTGACGATCG : 1922

Re-4 : 2380 \* 2400 \* 2420 \* 2440 \* 2460 \* : 2467  
BATG/VRe4-1 : CCAACGCCAACGGGACGACAATGCCCTCCGTGCAAGCCTTGTGGCAAAATGATGACATTGCTCGCTCCCTTGACTTGATCTTGGCAGCAACGATC : 2017

Re-4 : 2480 \* 2500 \* 2520 \* 2540 \* 2560 : 2562  
BATG/VRe4-1 : GTGCTTCGGCGCTCGTATGGGGAACCGACAACCGGTTCTTTGTTGACGGAGCAGGCGTTGCTGTTCGGTAGTTGCTGCTCGGTTGGCGGAA : 2112

Re-4 : \* 2580 \* 2600 \* 2620 \* 2640 \* 2660 : 2657  
BATG/VRe4-1 : AGCGTTGTCGGAATTGCTGCTCCGGTGGCGGCGCGGTAGGGAAGTGTGCTTTGGGACCTGCGACGGGACGCGGTTTGGCGTTGCGGG : 2207

Re-4 : \* 2680 \* 2700 \* 2720 \* 2740 \* : 2752  
BATG/VRe4-1 : GCTTCGGTACCCGGATGTTTTTACCGGTACATTGGCAAGCGGTTGACGCTTGGAAGTGTGCGGAACCTAGATGGTATCTCTTCCC : 2302

Re-4 : 2760 \* 2780 \* 2800 \* 2820 \* 2840 \* : 2847  
BATG/VRe4-1 : GCGGAGGAGTCCGTCCTTGGCGGAAATGCGTGACGCAAAAGCCTCGCGTGTCCCACTCCCAAGGAGCAAGATCGCTGGATGCGTCGCGCGG : 2397

Re-4 : 2860 \* 2880 \* 2900 \* 2920 \* 2940 : 2942  
BATG/VRe4-1 : CCTGGGTCGTCACGGTTGCAATGTCCGGTACGTCCCGGATGTCCGTTACGTCTCGTCATGTCGCCATCGTTGAAGGACGATGGGACTCC : 2492

Re-4 : \* 2960 \* 2980 \* 3000 \* 3020 \* 3040 : 3037  
BATG/VRe4-1 : TTCCGGAAGAGGGCAGTCTCTCTCCGGAACCGAAAGGAAGCGACGCCCGCAACGGTCAAGAATGGACGTCGGATTCTGTCGAATTTGTTTTCTC : 2587

Re-4 : \* 3060 \* 3080 \* 3100 \* 3120 \* : 3132  
BATG/VRe4-1 : CCTCTCCCCCAGTAGCCCTACTGGAATGGACAGGATGCCACCTCTAGTGGTGAAGAATGGACCGCGACTCGTCGGCTTGGTTTCGGCA : 2682

Re-4 : 3140 \* 3160 \* 3180 \* 3200 \* 3220 \* : 3227  
BATG/VRe4-1 : TTGGCGGAGTGGATGAATAAAGGAGCGGATGCTTACATTTTCTTCTTGCCTACAGTCAGCATCCGCTCGCGATTTTGATTTGGGGT : 2777

Re-4 : 3240 \* 3260 \* 3280 \* 3300 \* 3320 : 3322  
BATG/VRe4-1 : TTGAAAAGTATTGAACAGCATTAGCTAACGAAAAGAAATAGTGACAAATACAAATGACTCTCGTCTCTCTTGTGATGCGACAGGTATCCGGC : 2872

Re-4 : \* 3340 \* 3360 \* 3380 \* 3400 \* 3420 : 3417  
BATG/VRe4-1 : TTTTGGCCATTGACCTTTGTAAGATGACCGTCTAATTAATTGCAAGACTCGGTGAAAAACGAATAATCAAGACAGATATATTATCCGATC : 2967

Re-4 : \* 3440 \* 3460 \* 3480 \* 3500 \* : 3512  
BATG/VRe4-1 : AGACCAAGTGAACCGCCAACTTGGCCAAGCGAAGCGGCTTATTAAACGAACAGTCTTAATCGAACAGCAACTCGCGGGCGACACAACTGCG : 3062

Re-4 : 3520 \* 3540 \* 3560 \* 3580 \* 3600 \* : 3607  
BATG/VRe4-1 : TCATCGGCATCCACCATCCACACCAAGCCGTCGCATTCACAACAGAGAACTCCTCCGGCACAACCGGTTTCGCAATACGCAAGCATTAAT : 3157

Re-4 : 3620 \* 3640 \* 3660 \* 3680 \* 3700 : 3702  
BATG/VRe4-1 : CTCGGGAGATTGCTACCATGCTCCACAAAAGACTGAGTGTGATTGCCAGGATACGAAAAAGGCAACAGATTGTTTCCAAGAGTCATTGGGA : 3252

Re-4 : \* 3720 \* 3740 \* 3760 \* 3780 \* 3800 : 3797  
BATG/VRe4-1 : GGTACGCTCAGATTGCCCTGCGAGGATCCGATAAGGCTTTCAAGTCTCGATCGCTATCTCTCGGTGACGGCCGATCTCCGCATACAGAGAA : 3303

Re-4 : \* 3820 \* 3840 \* : 3855  
BATG/VRe4-1 : ATTTGAATCCAACGACTTGGAGCCCAAGCGCGACATTTTGGCGGGCCATATATCTG : -

## COMMUNICATIONS ET PUBLICATIONS

### Communications orales

**2013** D.H. Nguyen, D. Hermann, A. Caruso, J. Marchand, E. Tastard, S. Thiriet-Rupert, N. Casse, A. Morant-Manceau, "Un transposon de type *mariner* (Mariner-Like-Element) chez la microalgue marine *Amphora acutiuscula* (Bacillariophyta) : caractérisation et expression en conditions de stress thermique". **XVIII<sup>ème</sup> Congrès National sur les Eléments Transposables (CNET)**, 1 - 3 juillet, Montpellier – France

**2012** Nguyen D.H., A. Caruso, A. Morant-Manceau, D. Hermann, J. Rouault, S. Hiard, N. Casse. Eléments transposables de type *mariner* et leur expression en condition de stress thermique chez la diatomée marine *Amphora acutiuscula*. **2<sup>ème</sup> édition des journées scientifiques de l'école doctorale VENAM**, 25 et 26 Octobre à Nantes - France.

**2012** Concours du PRIX DU POSTER dans le cadre du Forum Jeunes Recherche 2012. Ecole Doctorale VENAM, le 22 Novembre, Le Mans - France

**2012** Nguyen D.H., Caruso A., Morant-Manceau A., Hermann D., Rouault J., Hiard S., Casse N., "Eléments transposables de type *mariner* chez la diatomée marine *Amphora acutiuscula*", 31<sup>ème</sup> **Colloque International** de l'Association des Diatomistes de Langue Française (ADLaF), 11-13 septembre, Le Mans, France

### Communications affichées

**2012** Nguyen D.H., Caruso A., Morant-Manceau A., Hermann D., Tastard E., Rouault J., Hiard S., Casse N., "Characterization and expression of Mariner-Like Transposons in the marine microalga *Amphora acutiuscula* (Bacillariophyceae)". **8th Asia-Pacific Conference on Algal Biotechnology (APCAB)**, 9 - 12 Juillet, Adelaide – Australia

**2012** Nguyen D.H., Caruso A., Morant-Manceau A., Hermann D., Rouault J., Hiard S., Casse N., "Characterization and expression of Mariner-Like Transposons in the marine microalga *Amphora acutiuscula* (Bacillariophyceae)". **International Congress on Transposable Elements (ICTE)**, 21 - 24 Avril 2012, Saint Malo – France

**2011** Hermann D., Nguyen H.D., Casse N., Jacquette B., Morant-Manceau A., Caruso A., "Expression des éléments transposables de type *mariner* en conditions de stress thermique et identification de leurs structures complètes chez la microalgue marine *Amphora acutiuscula*". **XVII<sup>ème</sup> Congrès National sur les Eléments Transposables (CNET)**, 4 - 6 juillet, Lyon – France

### Publications

**2014** Hermann D, Egue F, Tastard E, Nguyen DH, Casse N, Caruso A, Hiard S, Marchand J, Chénais B, Morant-Manceau A, Rouault J. Introduction to the vast world of transposable elements - what about the diatoms? **Diatoms Research**, 29 (1), 91-104

**2014** Nguyen DH, Hermann D., Caruso A., Tastard E., Marchand S., Rouault JD., Denis F., Rupert-Thiriet S., Casse N. and Morant-Manceau A. First evidence of Mariner-Like transposons in the genome of the marine microalga *Amphora acutiuscula*. **Protist**, 165, 730-744



## REVIEW

### An introduction to the vast world of transposable elements – what about the diatoms?

DOROTHÉE HERMANN<sup>1</sup>, FADOUMO EGUE<sup>1</sup>, EMMANUELLE TASTARD<sup>1</sup>, DUC-HUNG NGUYEN<sup>1</sup>, NATHALIE CASSE<sup>1</sup>, AURORE CARUSO<sup>1</sup>, SOPHIE HIARD<sup>1</sup>, JUSTINE MARCHAND<sup>1</sup>, BENOÎT CHÉNAIS<sup>1</sup>, ANNICK MORANT-MANCEAU<sup>1\*</sup> & JACQUES DÉRIC ROUAULT<sup>2</sup>

<sup>1</sup>PRES LUNAM, Université du Maine, Laboratoire Mer, Molécules, Santé (EA 2160), Le Mans, France;

<sup>2</sup>Laboratoire Evolution Génomes Spéciation, CNRS UPR 9034, Orsay, France

Transposable elements (TEs) are DNA sequences that are able to move and replicate within the genomes of virtually all organisms, ranging from animals to plants, and also including protozoans, bacteria and diatoms. Initially considered to be ‘junk DNA’ with no function, TEs are now thought to play a pivotal role in the evolution of genomes and organisms. This review is intended to provide non-specialists with an introduction to the world of TEs by summarizing the information available about these elements, ranging from their impact on genome structure to their potential involvement in the evolution of species. In this review, we propose to bring together what is known about the TEs and provide an overview of recent advances concerning TEs in diatoms.

**Keywords:** *transposable element, genome, diatom, evolution, stress response*

#### Introduction

Transposable elements (TEs) are mobile DNA sequences that are able to move within genomes. For a long time these elements were known as ‘junk DNA’, because they contain large numbers of repetitions and short sequences, and often congregate in non-coding regions of DNA, such as the heterochromatin. TEs display great diversity and are ubiquitous in living organisms, which led researchers to study their involvement in genomics. Improvements in sequencing techniques and data analysis are making it easier to identify these elements in a wide range of organisms. Current genomic developments are providing useful tools for investigating the activity of TEs and determining their impact on genome shape.

Here, we describe the characteristics of TEs and relate their history from their initial discovery to the latest findings in the ecological context, with examples taken from animals and plants, before going on to focus on TEs in diatoms, and what they can tell us about the genomes they inhabit. Diatoms are unicellular algae that display some interesting features, such as having small genomes and a single-cell structure, and which have successfully adapted to all aquatic environments. They provide good models for studying genomics and the physiology of photosynthetic organisms. They are also one of the most important components of phytoplankton, and are of huge ecological interest. By generating potential genome diversity and adaptation, TEs may play a key role in the worldwide distribution and

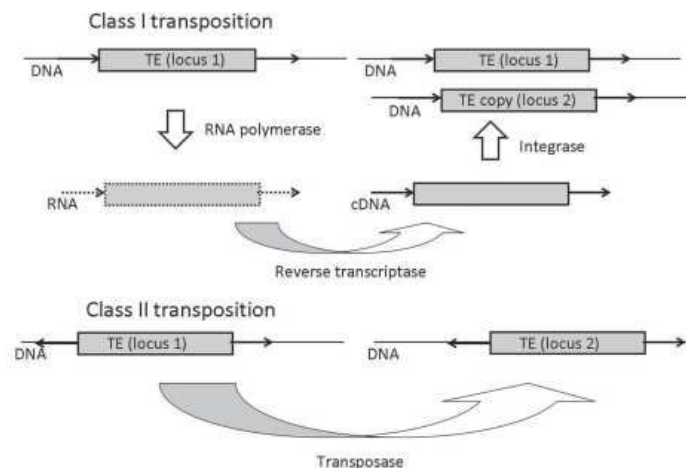
biodiversity of diatoms. Studying TEs can be expected to provide information about the evolution of diatoms.

#### From the discovery of TEs to the concept of the dynamic genome

For many years, the genome was thought to be an assembly of numerous genes, each with its own function and associated with a specific locus on a chromosome within a fixed structure. This overly narrow vision was completely revolutionized by the discovery of non-coding DNA and genetic instability. In the 1940s, the geneticist Barbara McClintock was working on maize diversity, and more particularly on how to improve maize varieties. Investigating the heritability of some characteristics, she observed particular alterations in the color of ears of corn, and demonstrated that these mutations were reversible, even though they were transmitted to the next generation (McClintock 1951). This conflicted with Mendel’s laws, and so McClintock introduced the concept of ‘controlling elements’ for the genetic loci that control gene function. She later found that these mutations were, in fact, induced by several controlling elements acting as activators and/or repressors and which are able to move within the chromosome. McClintock hypothesized that this movement, or transposition, might occur following genomic challenges such as double-strand breaks, and so she called them activation/dissociation elements (Ac/Ds; McClintock 1984). Unfortunately, McClintock

\*Corresponding author. E-mail: [annick.manceau@univ-lemans.fr](mailto:annick.manceau@univ-lemans.fr)

(Received 17 December 2012; accepted 6 December 2013)



**Fig. 1.** Schematic representation of the main transposition mode of class I retroelements and class II DNA TEs. Class I TEs: example of the LTR retrotransposons, which transpose by generating a copy of the original TE thanks to an RNA polymerase. Subsequently, the RNA copy is reverse-transcribed into complementary DNA (cDNA), and the cDNA is then inserted at another locus. Class II TEs: example of the TIR-transposons which transpose directly. The transposase recognizes the TIR and takes charge of the whole element, which is then inserted into the genome at another locus.

was far ahead of her time, and the scientific community was skeptical about her theory. The new vision of genomics implied by her hypothesis called for a complete overhaul of the existing narrow vision of the genome as a fixed entity. She finally gained credibility and was awarded the Nobel Prize in 1983. Since then, these ‘control elements’ have been extensively studied, and it is now established that genomes are flexible and able to adapt, in part thanks to these TEs. More than 29 000 studies on TEs have been published, and we now know that TEs are present in all genomes, that they sometimes make up a huge fraction of the genome (particularly of the biggest genomes, such as those of plants), that they can be involved in gene creation and regulation, and are usually involved in genome evolution. They are far from having revealed all their secrets.

#### The ID cards of TEs

TEs are DNA sequences that can be as short as 80 bp, but can reach >20 000 bp in length (Yeadon & Catchside 1995, Chopra *et al.* 1999, Kramerov & Vassetzky 2005, Macas & Neumann 2007, Wicker *et al.* 2007), and are characterized by their ability to move and/or replicate within the genome. This means that TE sequences can move from one DNA locus to another, which may be nearby or far away (Fig. 1). TEs are included in the repetitive DNA fraction that also contains endogenous retroviruses, tandem repeats and microsatellites. They can have diverse structures and, as such, no universal structural rule exists for their identification. The lack of general structural features makes them very difficult to identify. They are often

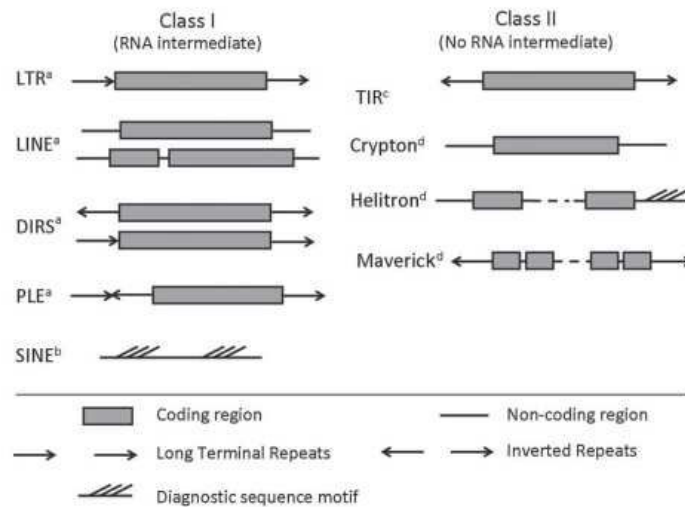
detected on the basis of their similarity to and/or homology with established TEs, hence TEs with novel structures or carrying numerous mutations are still difficult to identify and classify. TE diversity and the frequent discovery of new elements make their classification challenging.

#### Classification of TEs

TEs were originally divided into two classes depending on the nature of their transposition intermediate and main mechanisms of movement (Finnegan 1989, Capy *et al.* 1998; Fig. 1). The commonly accepted classification is currently that proposed for eukaryotic TEs by Wicker *et al.* (2007). This bases the first level of classification on the presence or absence of an RNA transposition intermediate, with subsequent classification levels being determined by the mode of transposition, the insertion mechanism that leads to target site duplication (TSD) at the end of the insertion process, their structural organization and sequence homologies.

- Class I TEs were designated ‘retrotransposons’ because their transposition machinery includes a reverse transcription process resulting in one or several additional copies of the original element. These copies are subsequently inserted at other loci (Fig. 1). This replicative mechanism is known as ‘copy-and-paste’, and can result in the presence of numerous copies of the original element in the host genome. This type of mechanism might lead to the rapid invasion of the genome by retrotransposons. Class I TEs consist of five orders (Fig. 2) containing a total of





**Fig. 2.** Schematic representation of the structures of the main orders of TEs. Classes of TEs are distinguished on the basis of the nature of the intermediates used during transposition and their insertion target sites. Different orders reflect the insertion mechanism and general gene organization. LTRs, long-terminal repeats; LINEs, long interspersed nuclear elements; DIRS, *Dictyostelium* intermediate repeat sequences; PLEs, *Penelope*-like elements; SINEs, short interspersed nuclear elements; TIRs, terminal inverted repeats. (a) Autonomous elements contain one or several open reading frames which encode several proteins, including a reverse transcriptase and (b) non-autonomous elements usually contain one ORF encoding a transposase, and may also contain one additional ORF and/or additional repeats; they may contain one or several ORFs coding different proteins but no transposase.

17 superfamilies, which are grouped according to the organization of their gene domain and the presence or absence of smaller direct or inverted repeats of non-coding domains: (1) five superfamilies (*Ty1-Copia*, *Ty3-Gypsy*, *Bel-Pao*, *Retrovirus*, *ERV*) make up the LTR order; (2) four superfamilies (*R2*, *RTE*, *Jockey*, *L1*) make up the LINE order; (3) three superfamilies (*DIRS*, *Ngaro*, *VIPER*) make up the DIRS order; (4) the *Penelope* superfamily alone constitutes the PLE order, and finally (5) three superfamilies (*tRNA*, *7SL*, *5S*) make up the SINEs.

- Class II TEs have been designated 'DNA transposons' because they transpose via a DNA intermediate. Their transposition involves excision of the element and its subsequent reinsertion at another locus in the genome. This mechanism, which is also described as 'conservative', is usually known as 'cut-and-paste' because it is the original element itself that is reinserted (Fig. 1). In theory, this conservative transposition is a less powerful genome-invading tool, but all duplication events help to increase the transposon population in a given genome and the transposition that occurs during chromosome replication. Class II TEs are not exclusively cut-and-paste TEs (subclass 1); some are TEs with a DNA intermediate that use a replicative transposition mechanism involving the displacement of only one DNA strand (subclass 2). Class II consists of four orders (Fig. 2). The TIR order contains nine superfamilies including

the prokaryote insertion sequences (IS) (*IS630-Tc1-mariner*, *hAT*, *Mutator*, *Merlin*, *Transib*, *P*, *Piggy-Bac*, *PIF-Harbinger* and *CACTA*) and all the TEs with TIR regions at each extremity. The *Crypton*, *Helitron* and *Maverick* orders are each composed of a single superfamily.

#### Can TEs move by themselves?

Although the distinguishing characteristic of TEs is their ability to move and replicate within genomes, some elements are either inactive (immobile remnants) or have lost the ability to move by themselves (non-autonomous elements), even though they still belong to the TE fraction of genomes. Remnants have lost their ability to move as a result of sequence deletion or mutation, and only part of the original element is still present in genomes. They certainly constitute the bulk of TEs in genomes, but they are also the most difficult to estimate because they may be so mutated or deleted that they are no longer recognizable. Non-autonomous elements are composed of TEs that use the machinery of other TEs to transpose and replicate. The TEs that help non-autonomous elements to move are known as 'master' elements (Feschotte & Mouchès 2000). A single master element can handle numerous non-autonomous elements and help them to invade the genome (Wessler et al. 1995, Feschotte & Mouchès 2000, Macas et al. 2005). Non-autonomous elements often display a similar structure and/or similar sequence homologies to their master

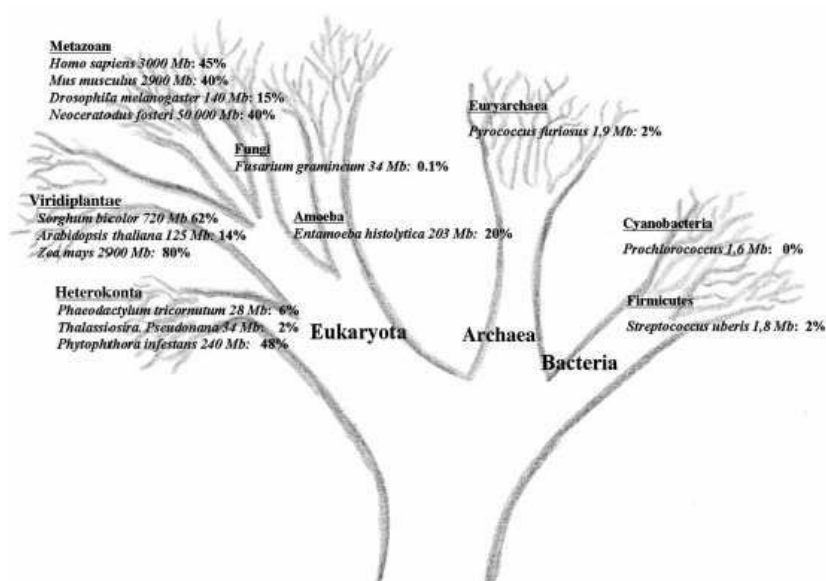


Fig. 3. Schematic overview of the percentages of TEs in the genomes of different organisms in the tree of life. The percentages of TEs in genomes were estimated by Landers *et al.* (2001), Chinwalla *et al.* (2002), Rocap *et al.* (2003), Armbrust *et al.* (2004), Clark *et al.* (2007), Filee *et al.* (2007), Merchant *et al.* (2007), Bowler *et al.* (2008), Lorenzi *et al.* (2008), Haas *et al.* (2009), Paterson *et al.* (2009), Schnable *et al.* (2009), Ward *et al.* (2009), Ma *et al.* (2010) and the *Arabidopsis* Genome Initiative (2000). Genome sizes are indicated in Megabases (Mb) next to the name of the organism.

element. In particular, they may harbor signal sequences that are recognized by the transposition enzymes of their master TE. Non-autonomous elements are probably the most widely distributed TEs in some genomes (Pereira *et al.* 1986, Jiang *et al.* 2004, Coates *et al.* 2012). This raises the following questions: how do TEs invade genomes, which genomes do they invade and what is their life cycle?

#### TEs as invaders of genomes

To date, almost all genomes in which TEs have been searched for, whether of marine (Derelle *et al.* 2006, Bui *et al.* 2008, Maumus *et al.* 2009, Piednoël & Bonnivard 2009) or terrestrial species, ranging from human beings (Smit 1996) to prokaryotes (Leclercq *et al.* 2011) and including plants (Bureau & Wessler 1994, Park *et al.* 2011), vertebrates (Robertson 1993, Biémont *et al.* 1997, Sinzelle *et al.* 2005, Casola *et al.* 2008), invertebrates (Arkhipova & Meselson 2005), metazoans and fungi (Kempken & Windhofer 2001, Daboussi & Capi 2003), have been shown to have been invaded by these elements (Fig. 3). TEs actually make up most of the genome in many species (Fig. 3). For instance, in maize and rice, TEs account for c. 60–80% of the genome. These levels can vary considerably even between closely related species. In rice, TEs account for 25–70% of the genome, depending on the species or cultivar studied (Zuccolo *et al.* 2007). Large genomes are

mainly invaded by TEs (Fig. 3), and so it is assumed that proliferation is a possible cause of an increase in genome size (Haas *et al.* 2009, Grover & Wendel 2010, Chénais *et al.* 2012, Muñoz-Diez *et al.* 2012). TEs make up <10% of small genomes, such as those of yeast or the microalga *Chlamydomonas reinhardtii* (Biémont & Vieira 2006, Merchant *et al.* 2007, Chénais *et al.* 2012). However, the relationship between genome size and TE percentage is not linear (Kidwell 2002, Piednoël *et al.* 2012). It has been posited, therefore, that genomes in which TEs are less numerous are either those that have undergone major genomic rearrangements or are of species with large populations and rapid evolution, facilitating the elimination of TEs, as in the case of prokaryotes (Treangen *et al.* 2009, Cerveau *et al.* 2011). These percentages are only estimations, because large fractions may remain undetected among the unclassified/unidentified DNA sequences obtained by genome sequencing (Burgess 2012).

Genomes exhibit differing proportions of a wide diversity of TEs (Hua-Van *et al.* 2005, Feschotte & Pritham 2007). In general, class I elements are predominant in large genomes, such as those of plants, mammals and algal species (Landers *et al.* 2001, Merchant *et al.* 2007, Vitte *et al.* 2007, Joseph *et al.* 2009, Maumus *et al.* 2009, Park *et al.* 2012), whereas bat and fungal genomes seem to have been invaded mainly by class II elements (Kempken & Kück 1998, Feschotte & Pritham 2007, Ray *et al.* 2008, Ogasawara *et al.* 2009).



### The life cycle of TEs in genomes

According to Kidwell & Lisch (2001), the TE life cycle consists of three stages: (1) invasion, (2) maturation and (3) senescence. First, an active and autonomous TE (i.e., one able to transpose by its own) inserts into the genome, probably following a horizontal gene transfer (HGT) event. After insertion, the TE replicates in the genome until it induces deleterious effects by altering gene function and thus impairs host fitness. Second, when deleterious effects are experienced by the host, selection processes come into play and limit proliferation of the TE. These regulatory mechanisms can, for instance, involve repressing transposition (by factors that bind to the gene and control its transcription), density-dependent autoregulation (an increase in the transposition enzyme can lead to negative feedback that reduces TE activity; Machida & Machida 1989, Zerbib et al. 1990, Zhang et al. 2001, Nagy & Chandler 2004, Jalliet et al. 2012, Claeys Bouuaert et al. 2013) or epigenetic processes. The latter consist of two types of silencing: DNA methylation or post-transcriptional silencing as a result of RNA interference and histone modifications (Slotkin & Martienssen 2007, Blumestiel 2011). These mechanisms silence TE expression at the transcriptional or post-transcriptional stage, respectively. The third and final phase is marked by a drastic decrease in the copy number of the original TE, until it is eventually purged from the host genome altogether. Some of these processes might have been responsible for the total elimination of TEs from the genomes of certain species, such as the obligate intracellular prokaryotes, which have not therefore experienced reinvasion by HGT (Wright & Finnegan 2001, Bordenstein & Reznikoff 2005, Cordeaux 2009).

Counter-selection (Le Rouzic & Capy 2005), stochastic loss or genetic drift can all lead to the elimination of certain TEs from the genome. It has been suggested that these mechanisms are responsible for the sporadic distribution of the *mariner* family of transposons in *Drosophila* populations (Lohe et al. 1995, Wang et al. 2005). A subsequent invasion phase is also possible as a result of HGT events from other organisms, such as bacteria or parasites. Although it is difficult to assess these lateral transmissions (Diao et al. 2006, Fortune et al. 2008), they are thought to have occurred repeatedly (Robertson & Lampe 1995, Silva & Kidwell 2004, Roulin et al. 2008, Gilbert et al. 2010, Wallau et al. 2011). Other mechanisms may have been responsible for a new active element or a new invasion during the first stage. For instance, the reactivation and/or re-colonization of an element being active by the recombination between of existing copies is also another credible scenario, and reactivation by stressful conditions can be also assumed to occur, as the findings of some studies support both these hypotheses (Sharma et al. 2008, Alzohairy et al. 2012). Another possibility is that some mutated copies may escape from the regulatory pathways and lead to new bursts of transpositions, although this is difficult to demonstrate. The link between the beginning and the end of the TE

life cycle remains blurred, and it is difficult to propose any general rules, since several different processes can occur simultaneously, and have joint impacts on the evolution of TE populations.

### Impact of TEs on genomes

#### *TE activity can mess things up*

Because of their activity and massive presence in genomes, TEs can also be viewed as parasitic sequences inducing potentially deleterious effects. At first glance, the most obvious impact is that on genome size (Chénais et al. 2012). For instance, some genomes, such as those of plants, have been so massively invaded that TEs now make up most of the genome. It is obvious that harboring so many elements must have affected genome structure by facilitating chromosomal sequence rearrangements and disrupting gene expression (Bourque 2009, Chénais et al. 2012). The latter has the most obvious impact, and can lead to phenotype shifts such as variegation (Hashida et al. 2003, 2006, Takahashi et al. 2012). This is the phenomenon that led to the original discovery of the first TEs: *Ds-Ac* in maize (McClintock 1951) and *mariner* in *Drosophila* (Hartl 2001), which have become some of the most extensively studied. Another example of gene disruption, this time with deleterious consequences, is that of hemophilia, a disease caused by insertion of the *L1* TE into the gene on the human X chromosome that encodes blood factor VIII (Dombroski et al. 1991). However, TEs do not have only deleterious effects; some can play positive roles in genomes.

#### *Are TEs beneficial?*

It has been posited that possible TE interactions with their host genome can be seen as a continuum extending from the parasitic to the symbiotic (Kidwell & Lisch 2001). Indeed, TEs are quite often studied because of their potential impact on genome evolution. TEs can be a source of gene regulatory tools that influence the expression of neighboring genes, and several examples of promoters, enhancers or *cis*-regulatory sequences derived from TEs have been described (Sinzelle et al. 2009, Chénais et al. 2012). TEs can also be a source of genetic innovation through the 'domestication' phenomenon, in which the TE sequence is integrated into the genome and subsequently plays a direct role in a cellular function (Sinzelle et al. 2009, Chénais et al. 2012). An instance of this is the V(D)J vertebrate immune system, which is derived from the TE *Transib* (Kapitonov & Jurka 2005). Therefore, TEs can be involved in various key roles, such as: (1) reorganization of the genome after polyploidization (Parisod et al. 2010), (2) gene regulation by combining with transcription factors (Testori et al. 2012), (3) the maintenance of chromosome integrity during embryo division (Khurana et al. 2010), or (4) in X-chromosome inactivation to promote male gene expression in late spermatogenesis (Lyon 2000, Hense et al. 2007). TEs may even have been



involved in the evolution of sex chromosomes and in the evolution of life (Charlesworth *et al.* 2005, Bachtrog 2006). Thus, TE insertion into the genome may have positive effects on its fitness, as seen in the domesticated TEs.

#### *TEs and stress responses*

TE activities have often been shown to be tightly linked to various biotic and abiotic stressful conditions, such as temperature changes, UV treatment, nutrient starvation or pathogen attacks (Mhiri *et al.* 1997, Ilves *et al.* 2001, Kimura *et al.* 2001a, b, Bouvet *et al.* 2008, Maumus *et al.* 2009, Ogasawara *et al.* 2009, Rakocovic *et al.* 2009). Some of these studies have revealed that TE expression or activity was induced concurrently with the induction of defense molecules (Mhiri *et al.* 1997, Takeda *et al.* 1999, Kimura *et al.* 2001b), leading to the suggestion that these TEs may be implicated in either the global stress response or an adaptive TE response to certain specific stressful conditions. Moreover, it has been demonstrated that TEs and stress-responsive genes share regulatory sequences that might activate or enhance TE expression, especially as part of the stress response (Strand & McDonald 1985, Grandbastien *et al.* 1997, Bouvet *et al.* 2008). For instance, the *Cyp6g1* gene, which is associated with the *Accord* retrotransposon, forms a gene complex that is overexpressed in *Drosophila* exposed to insecticides, leading to internal physiological detoxification (Chung *et al.* 2007). This gene combination has been found in natural populations of *Drosophila* from around the world, suggesting that it has been selected for in the population as a useful adaptation process. Stress-related TE activation can also result from a shift in methylation status that corresponds to the abolition of silencing of the TE gene structure, as demonstrated for the *Tam3* transposon in snapdragon. *Tam3* transposition is induced by hypomethylation of the transposon resulting from changes in temperature (Hashida *et al.* 2003, 2006). This does not reflect an adaptive process, but rather a concurrent effect of genetic modification resulting from the environmental change.

#### **Do TEs mediate environmental adaptation?**

The differing distributions of stress-responsive TEs in natural populations could be an argument in favor of the theory that TE activity is a key factor in genomic adaptation to environmental change (Oliver & Greene 2012), as first formally proposed by Barbara McClintock (1984). If TEs are to be implicated in environmental adaptation, it is critical to find out whether TE expression or activity can be beneficial in environmental adaptation processes. Studying barley plants (*Hordeum spontaneum*) from 'Evolution Canyon', Israel, where plants are subject to markedly different microclimates, Kalendar *et al.* (2000) noticed different class II TE *Bare-1* insertion profiles in barley taken from different sites on the canyon slopes. The two sides of the canyon are

exposed to differing solar irradiation and humidity levels and the invasion of these barley plants by *Bare-1* increases with their height on the canyon slopes, the plants at the top being in the driest part. This finding correlates with the overall decrease in species diversity with increasing height on the canyon slope, and suggests that *Bare-1* TEs have induced genome restructuring. One may suppose that activation of these TEs might be associated with adaptation of the barley to growing in 'Evolution Canyon'. As has been strongly suggested in the *Drosophila* genus, TEs may mediate adaptation to climate, since different profiles of TE insertion, such as that of *mariner* elements, are found in populations from different geographical locations (Chakrani *et al.* 1993, Giraud & Capy 1996). In a recent study, TE insertions and nearby genes were correlated to phenotype adaptation to temperate climate in a study of *Drosophila* populations that have adapted to a temperate climate (populations collected from temperate areas of Australia or North America), which contrasts with the tropical origin of the *Drosophila* genus (South Africa). Gonzalez *et al.* (2008, 2010) suggest that TEs located near sets of genes (e.g., those that regulate the circadian rhythm) and the insertion profile of different TEs in genes involved in fitness (such as the those governing odor-guide behavior related to the use of different food resources) are correlated with climate adaptation. These findings provide clues about the potential involvement of TEs in natural adaptive processes.

#### *Involvement of TEs in speciation mechanisms*

The involvement of TEs in speciation mechanisms remains controversial within the scientific community, although a few examples definitely suggest that TEs might be involved in speciation events. Recent studies, particularly those carried out in plants, suggest that TEs probably are involved in species evolution (Tian *et al.* 2011, Levy 2013). Because insertions and bursts of transposition (i.e., rapid and widespread TE activity) can lead to severe mutations and, sometimes, lethal effects, accumulations of TEs could impact on organism fitness (Pasyukova *et al.* 2004). The invasion of allopatric populations of *Drosophila* by TEs can reduce the fitness of their hybrids, and this hybrid dysgenesis may be the first step toward speciation (Rubin *et al.* 1982, Rose & Doolittle 1983). Recent studies have correlated these transposition 'bursts' with host lineage divergences in salmonids, bats and *Drosophila*. For instance, in the bat genus *Myotis*, the most diverse mammalian genus, repeated invasion of the *Myotis* genomes by transposons has been correlated with the wide genetic diversity associated with species divergence (Ray *et al.* 2008). Another example is sterility in *Drosophila* caused by TE transposition in the third chromosome of some *D. simulans*-*D. melanogaster* hybrids, leading to the reproductive isolation that constitutes the first step in the speciation process. Taken together, these findings suggest that TE accumulation may have been



a major factor in the emergence of new species, because it can facilitate genetic diversity (Masly et al. 2006, De Boer et al. 2007, Feschotte & Pritham 2007, Oliver & Greene 2009, 2012, Crespi & Nosil 2012).

#### Why look at TEs in diatoms ...

Diatoms are unicellular microalgae encased within a silica shell. They belong to the Stramenopile taxon, and constitute the most diversified group of photosynthetic eukaryotes, including c. 100 000 species. They have been reported in all aquatic environments ranging from marine to freshwater (Round et al. 1990), in which they constitute a key factor in biogeochemical cycles, such as the silicon and carbon cycles where they are responsible for c. 25% of global carbon fixation (Falkowski et al. 1998, Granum et al. 2005) and for nearly 40% of marine organic carbon production (Nelson et al. 1995). Diatoms are usually divided into two main groups: the centric diatoms that exhibit radial symmetry, and the pennate diatoms that exhibit bilateral symmetry. Following improvements in sequencing methods, many researchers decided to pick out two diatoms as models for studying diatom genomics and algal metabolism. To date, four marine diatom genomes have been sequenced and are available: *Phaeodactylum tricornutum* (pennate), *Thalassiosira pseudonana* (centric), *Fragilariopsis cylindrus* (pennate) and *Pseudo-nitzschia multiseries* (pennate) (Scala et al. 2002, Armbrust et al. 2004, Lopez et al. 2005, Montsant et al. 2005, Bowler et al. 2008), and at least eight others are either currently being sequenced or are about to be so (Tirichine & Bowler 2011, EMBO workshop communications). Recent articles have highlighted the critical need to identify TEs in order to assemble the genome correctly (Yandell & Ence 2012). It is worth focusing on the identification of these elements in diatom genomes even though they challenge genome annotators. The role of TEs in diatom genomes and in their evolution may be investigated to: (1) make it easier to annotate diatom genomes; (2) use TEs as molecular markers in phylogenies or in population studies, as has been done in other organisms (Giraud et al. 1997, Monden et al. 2009, Meyer et al. 2012, Shirasawa et al. 2012, Suh et al. 2012); (3) engineer molecular tools for microalgal transformation; and finally, (4) provide a model to evaluate their possible role in the great diversity and adaptation capacities of diatoms worldwide.

#### Diversity of diatom TEs

Diatom genome sequencing projects also initiated the TE content analysis of *P. tricornutum* (pennate), *T. pseudonana* (centric) has been estimated (Armbrust et al. 2004, Bowler et al. 2008, Maumus et al. 2009), and TE component analyses of *Ps. multiseries* and *F. cylindrus* are still in progress (F. Maumus & M. Parker pers. comm., F. Maumus & T. Mock pers. comm.). It has been estimated that TEs make up 6.4% and 1.9% of the *P. tricornutum* and

*T. pseudonana* genomes, respectively (Maumus et al. 2009), which suggest low levels of invasion by TEs in diatom genomes. This is consistent with previous observations that most organisms with small genomes (27.4 Mbp for *P. tricornutum* and 32.4 Mbp for *T. pseudonana*) contain small amounts of TEs. For the moment, only the TE contents of diatom species with small genomes are available. Analysis of the *Ps. multiseries* genome (218, 73 Mbp) or other bigger diatom genomes could extend our knowledge about TE diversity and dynamics. The TE fractions of both the diatom genomes that have been sequenced consist predominantly of Class I TEs. Indeed, almost 90% of the TEs found in *P. tricornutum* and 58% of those found in *T. pseudonana* belong to the class I TEs, particularly the LTR retrotransposons superfamily (Maumus et al. 2009) (Table 1). Representatives of the main TE superfamilies have been found in both *P. tricornutum* and *T. pseudonana*. For example, TEs belonging to the *Tc1-mariner* superfamily are present in *P. tricornutum* and in other pennate and centric species (Hermann 2011). Although we did not succeed in identifying them in the *T. pseudonana* genome, we did observe the presence of *Tc1-mariner* transposons in another centric diatom, *Thalassiosira weissflogii* (Hermann 2011). Studying TEs in diatoms has shed light on the history of the TEs, but can TEs shed light on the evolution of diatoms?

#### What TEs reveal about the structure and evolution of diatom genomes

Phylogenetic analysis of *Ty1-Copia* conserved reverse transcriptase domains from *P. tricornutum* and *T. pseudonana* genomes reveals the existence of seven groups designated CoDi1 to CoDi7 (*Ty1-Copia*-like elements from diatoms). The CoDi1, CoDi2, CoDi3 and CoDi7 groups are specific to *P. tricornutum*, whereas groups CoDi4, CoDi5 and CoDi6 are common to both these diatoms. The groups can be divided into three different lineages: (1) CoDiI (including CoDi1, -2, -3, -4 and -7) mostly found in *P. tricornutum*; (2) CoDiII (CoDi5 group) composed of elements belonging to both *P. tricornutum* and *T. pseudonana*; and (3) the *Copia* lineage (CoDi6), which includes elements from both diatoms and is closely related to *Ty1-Copia* elements from other organisms. This *Copia* lineage is the only one that contains elements very similar to those found in plants and animals, the others being specific to diatoms. The authors favor the hypothesis that HGT explains the presence of CoDi4 sequences in *T. pseudonana*, because they could not detect any remnants of other CoDi1, -2 and -3 TEs in centric diatom genomes. Global phylogenetic analysis of *Ty1-Copia* elements related to similar elements from other ecologically relevant micro-planktonic organisms, reveals that different marine organisms share some lineages of the *Ty1-Copia* retrotransposon. Moreover, CoDi6 belongs to a TE lineage that is phylogenetically related to *Copia* elements from mosquito and rice. This suggests either that



**Table 1.** Estimated proportions of TEs in diatom genomes.

		Diatom species			
		<i>Phaeodactylum tricornutum</i>	<i>Thalassiosira pseudonana</i>	<i>Pseudonitzschia multiseries</i>	<i>Pseudonitzschia multistriata</i>
TE order	TE superfamily				
LTR	<i>Ty1-Copia</i>	Psc > 85%/Exp	Psc > 56%	Psc n.a.	Psc n.a.
LTR	<i>Ty3-Gypsy</i>	Psc < 1%	Psc < 7%	Psc n.a.	Psc n.a.
LINE	<i>RTE</i>	X	Psc < 5%		
TIR	<i>Harbinger</i>	X	Psc < 25%		
TIR	<i>PiggyBac</i>	Psc < 5%	Psc < 2%		
TIR	<i>Mutator</i>	Psc < 6%	Psc < 5%		
TIR	Subclass I TEs	Psc < 3%	Psc n.a.		
TIR	<i>Tc1-mariner</i> -like	Psc n.a. Exp	X	Psc n.a.	

Note: The TE orders and superfamilies that have been identified in diatom genomes are summarized. The presence (Psc) or apparent absence (X) of each TE superfamily identified in each species is reported. When the presence of a superfamily has been detected, the estimated percentage relative to the total TE abundance and their expression (Exp) are both reported. n.a.: data not available. The percentages of TEs in diatom genomes were estimated from recent studies by Armbrust *et al.* (2004), Bowler *et al.* (2008), Maumus *et al.* (2009), Hermann (2011).

this CoDi6 lineage was integrated into diatom genomes following HGT events or that *Ty1-Copia* were present in the root of the eukaryote lineage, and have subsequently experienced different divergence events. Kapitonov & Jurka (2004) also propose that HGT events might explain the presence of *Harbinger* transposons (belonging to the TIR order) in *T. pseudonana*. This explanation was based on the fact that *Harbinger* transposons from rice and from *Arabidopsis thaliana* are phylogenetically closer to the diatom *Harbinger* elements than to *Harbinger* from closely related plant species. This adds credibility to the suggestion that HGT is probably more common in the evolutionary history of diatoms than in that of other organisms, as proposed by Bowler *et al.* 2008. In view of the distribution of TEs in both pennate and centric diatoms and their impact on gene regulation and genetic diversity, Bowler *et al.* (2010) suggested that TE might have been one of the forces behind diversification. These authors also suggest that TEs may have been responsible for diatom genome expansion, which is thought to have happened in plants (Vitte & Panaud 2005, Piegu *et al.* 2006).

#### TE expression in diatoms: standard versus stress conditions

Recent studies have reported mRNA expression in the TE fraction of diatom genomes (Table 1), which suggests that TEs have recently been active.

Norden-Krichmar *et al.* (2011) reported high levels of expression of DNA transposons and retrotransposons in the *T. pseudonana* cDNA library. In our laboratory, we were able to demonstrate the presence of *Tc1-mariner* mRNA transcripts in three diatom species, *P. tricornutum*, *Amphora coffeaeformis* and *Amphora acutiuscula*

(Hermann 2011, Nguyen *et al.* in prep.). In order to investigate TE expression, Maumus *et al.* (2009), analyzed the transcript abundance of the *Ty1-Copia* CoDi elements in expressed sequence tag (EST) databases established using different standard and stressed cultures of *P. tricornutum* and *T. pseudonana* (Maheswari *et al.* 2009) and revealed the expression of each group of diatom *Ty1-Copia* TEs under one or another of the 16 culture conditions tested for *P. tricornutum*. Similarly, in *T. pseudonana*, at least one of the six conditions tested to create the EST library revealed the expression of CoDi elements. Notably, in *P. tricornutum*, they observed a major increase in EST abundance corresponding to two retrotransposons: *Blackbeard* (which belongs to the CoDi 4 group) and *Surcouf* (which belongs to the CoDi 3 group) under conditions of nitrate starvation or toxic reactive-DD-aldehyde addition (decadienal aldehydes). The specific induction of *Blackbeard* and *Surcouf* expression under stressful conditions suggests that the transposition of these Class I elements might also have been activated and then contributed to the overall response/adaptation to stress. Both these elements were investigated further, and their differing insertion profiles in several populations of *P. tricornutum* suggest that they are active. *Blackbeard* expression was also investigated by Oliver *et al.* (2010) in a strain of *P. tricornutum* other than the one that has been sequenced. These authors recreated some of the conditions characteristic of algal blooms (nitrogen-depleted medium and high cell density), and investigated the expression of *Blackbeard* in a *P. tricornutum* strain (CAMP630) other than the one sequenced (CCMP2561), which had been used in previous studies (Bowler *et al.* 2008, Maumus *et al.* 2009). They observed a positive correlation between the increases in *Blackbeard* transcripts in cultures in nitrogen-depleted media, suggesting that nitrogen induces TE expression in diatoms. This



finding also suggests that cell density rather than stress is the main factor leading to enhanced expression of *Blackbeard*. Once the cell density factor had been integrated into their statistical analysis, Olivier et al. 2010 found a positive correlation between the enhancement of *Blackbeard* expression and the increase in the number of cells. Many studies have reported the activation or the enhancement of TE expression under conditions of abiotic or biotic stress in plants, animals, bacteria or fungi (for a review see: Chénais et al. 2012). Some studies have identified correlations between the induction of TEs and stress-responsive genes, suggesting that TE induction might be involved in the stress response of organisms (Grandbastien et al. 1997, Capy et al. 2000, Tapia et al. 2005).

In our laboratory, we are currently searching for the expression of *Blackbeard* and *Surcouf* retrotransposons in three marine diatoms including *P. tricornutum*. We have also shown that the *Tc1-mariner* superfamily is expressed in diatom genomes (Hermann 2011, Nguyen et al. in prep.).

#### *Diatom TEs and their potential ecological impact*

Maumus et al. (2009) investigated the causal correlation between stress and the overexpression of *Blackbeard* in *P. tricornutum*, and demonstrated that the LTR sequences of *Blackbeard* contain a *cis*-regulatory component responsible for induction of the expression of this TE in response to nitrate starvation. It has often been reported in the literature that TEs can exhibit regulatory sequences that are very similar to those regulating stress-responsive genes or transcription factors, especially in the *Ty1-Copia* superfamily of retrotransposons (Strand & McDonald, Tapia et al. 2005, Mhiri et al. 1997, Takeda et al. 1999, Beguiristain et al. 2001). These similarities might lead to the induction of TEs during the stress response. *Blackbeard* and *Surcouf* TEs might therefore act as environmental sensors in the response of diatoms to stress, something reinforced by the significance of the ecological relevance of both nitrogen- and DD-aldehyde-related stressful conditions. Indeed, as noted by Maumus et al. (2009), nitrate-replete conditions have a huge impact on the diatom population under natural conditions (Smetacek 1999) and unsaturated aldehydes are known to be intercellular signals that can influence cell death in diatoms (Casotti et al. 2005). Moreover, the results of Maumus et al. (2009) suggest that *Blackbeard* expression is probably under epigenetic control, which they point out, might suggest a direct link between environmental stress and chromatin modeling in the processes of adaptation to environmental changes. Norden-Krichmar et al. (2011) revealed that TEs from *T. pseudonana* are probably under tight epigenetic control that could mirror the regulation of TE expression by some environmental conditions. These studies show that diatom TE expression is probably associated with environmental and ecological changes. Oliver et al. (2010) may have revealed the factors that influence

diatom retrotransposon expression in natural environments by recreating some of the natural changes that occur under bloom conditions. This is particularly interesting, because it would make it possible to forecast TE expression at the population level. Oliver et al. concluded that *Blackbeard* expression is not correlated with physiological stress, but suggest that TEs are involved in the stress-sensing survey or in cell signaling that allows diatoms to adapt to changing environmental conditions. It is not possible to determine the impact of these TEs in diatom detoxification processes or in stress-signaling pathways because they are not correlated with any nominal cell function. However, it is clearly crucial to find out whether this correlation is a reality in order to see how TEs fit into the picture of how the whole cell responds to environmental conditions.

These results focused on particular TE superfamilies in two diatom species, but it is essential to find out whether TEs really do play a role in diatom responses to environmental changes and in diatom metabolism. If so, one may wonder whether this underpins the particular success of diatoms, since TEs might have induced the genetic diversity that has allowed diatoms to adapt successfully to so many environments, as suggested by Bowler et al. (2008) and Maumus et al. (2009).

#### Conclusion

TEs are now considered to be a major component of the genomes of living organisms. Their ability to move and replicate in genomes, their number and their involvement in numerous processes such as genome rearrangements, resistance to stress or adaptation to environmental conditions, mean that they are now generally viewed as key factors in the evolution of living organisms. To date, TEs have been studied in only a few diatom species, and only three TE superfamilies have been investigated. Because their ability to adapt has made diatoms one of the most successful components of the phytoplankton, due to and in view of the phylogenetic position of the diatoms in the eukaryote lineage, it is essential to investigate further the presence, diversity and role of TEs in diatoms. This will entail: (1) finding out more about TE superfamilies in diatoms; (2) finding out whether diatom TE expression/activities are linked to the stress resistance and fitness of diatoms and whether TEs do indeed contribute to the success of diatoms in all aquatic environments; (3) clarifying the activity of elements belonging to two different TE superfamilies, including class I and class II elements, in order to confirm whether TEs can be viewed as good candidates for mutagenesis (Sasakura et al. 2008, Delauriere et al. 2009, Palazzoli et al. 2010); and (4) continuing the search for autonomous and active elements in diatoms in order to explore the biotechnological potential of diatom TEs. They may have activities that could be exploited to create mutants



or catalyze the transformation of the genomes of diatoms or other marine organisms.

### Acknowledgements

This work was supported by the European program 570 GIAVAP (FP7-KBBE-2010-4). We would like to thank M. Ghosh for revising the English text; our especial thanks go to the two reviewers for their helpful comments that have enabled us to improve this manuscript.

### References

- ALZOHAIY A.M., YOUSEF M.A., EDRIS S., KERTI B., GYULAI G. & BAHIELDIN A. 2012. Detection of LTR Retrotransposons Reactivation induced by *in vitro* Environmental Stresses in Barley (*Hordeum vulgare*) via RT-qPCR. *Life Science Journal* 9: 5019–5026.
- ARKHIPOVA I.R. & MESELSON M. 2005. Diverse DNA transposons in rotifers of the class Bdelloidea. *Proceedings of the National Academy of Sciences* 102: 11781–11786.
- ARMBRUST E.V., BERGES J.A., BOWLER C., GREEN B.R., MARTINEZ D., PUTNAM N.H., ZHOU S., ALLEN A.E., APT K.E., BECHNER M., BRZEZINSKI M.A., CHAAL B.K., CHIOVITI A., DAVIS A.K., DEMAREST M.S., et al. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306: 79–86.
- BACHTROG D. 2006. A dynamic view of sex chromosome evolution. *Current Opinion in Genetics and Development* 16: 578–585.
- BEGUIRISTAIN T., GRANDBASTIEN M.A., PUIGDOMÈNECH P. & CASACUBERTA J.M. 2001. Three Tnt1 subfamilies show different stress-associated patterns of expression in tobacco. Consequences for retrotransposon control and evolution in plants. *Plant Physiology* 127: 212–221.
- BIÉMONT C. & VIEIRA C. 2006. Genetics: Junk DNA as an evolutionary force. *Nature* 443: 521–524.
- BIÉMONT C., TSITRONE A., VIEIRA C. & HOOGLAND C. 1997. Transposable element distribution in *Drosophila*. *Genetics* 147: 1997–1999.
- BLUMESTIEL J.P. 2011. Evolutionary dynamics of transposable elements in a small RNA world. *Trends in Genetics* 27: 1, 23–31.
- BORDENSTEIN S.R. & REZNIKOFF W.S. 2005. Mobile DNA in obligate intracellular bacteria. *Nature Reviews Microbiology* 3: 688–699.
- BOURQUE G. 2009. Transposable elements in gene regulation and in the evolution of vertebrate genomes. *Current Opinion in Genetic Developments* 19: 607–612.
- BOUVET G.F., JACOBI V., PLOURDE K.V. & BERNIER L. 2008. Stress-induced mobility of *OPHI1* and *OPHI2*, DNA transposons of the Dutch elm disease fungi. *Fungal Genetics and Biology* 45: 565–578.
- BOWLER C., ALLEN A.E., BADGER J.H., GRIMWOOD J., JABBARI K., KUO A., MAHESWARI U., MARTENS C., MAUMUS F., OTILLAR R.P., RAYKO E., SALAMOV A., VANDEPOELE K., BESZTERI B., GRUBER A., HEIJDE M., KATINKA M., MOCK T., VALENTIN K., VERRET F., BERGES J.A., BROWNLEE C., CADORET J.-P., et al. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456: 239–244.
- BOWLER C., VARDI A. & ALLEN A.E. 2010. Oceanographic and biochemical insights from diatom genomes. *Annual Reviews of Marine Sciences* 2: 333–365.
- BUI Q.T., CASSE N., LEIGNEL V., NICOLAS V. & CHÉNAIS B. 2008. Widespread occurrence of *mariner* transposons in coastal crabs. *Molecular Phylogenetics and Evolution* 47: 1181–1189.
- BUREAU T.E. & WESSLER S.R. 1994. *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. *Plant Cell* 6: 907–916.
- BURGESS D.J. 2012. Genomics: Repetitive elements underestimated? *Nature Reviews Genetics* 13: 74.
- CHARLESWORTH D., CHARLESWORTH B. & MARAIS, G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95: 2, 118–128.
- CAPY P., BAZIN C., HIGUET D. & LANGIN T. 1998. *Dynamics and evolution of transposable elements*. Landes Company, Austin. 197 p.
- CAPY P., GASPERI G., BIÉMONT C. & BAZIN C. 2000. Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85: 101–106.
- CASOLA C., HUCKS D. & FESCHOTTE C. 2008. Convergent domestication of *Pogo*-like transposases into centromere-binding proteins in fission yeast and mammals. *Molecular Biology and Evolution* 25: 29–41.
- CASOTTI R., MAZZA S., BRUNET C., VANTREPOTTE V., IANORA A. & MIRALTO A. 2005. Growth inhibition and toxicity of the diatom aldehyde 2-trans, 4-trans decadienal on *Thalassiosira weissflogii* (Bacillariophyceae). *Journal of Phycology* 41(1): 7–20.
- CERVEAU N., LECLERCQ S., BOUCHON D. & CORDAUX R. 2011. Evolutionary dynamics and genomic impact of prokaryote transposable elements. In: *Evolutionary biology – concepts, biodiversity, macroevolution and genome evolution* (Ed. by P. Pontarotti). Springer, Berlin. 345 pp.
- CHAKRANI F., CAPY P. & DAVID J. 1993. Developmental temperature and somatic excision rate of *mariner* transposable element in three natural populations of *Drosophila simulans*. *Genetics, Selection, Evolution* 25: 121–132.
- CHÉNAIS B., CARUSO A., HIARD S. & CASSE N. 2012. The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. *Gene* 1: 7–15.
- CHINWALLA A.T., COOK L., DELEHAUNTY K.D., FEWELL G.A., FULTON L.A., FULTON R.S., GRAVES T.A., HILLIER L.W., MARDIS E.R., MCPHERSON J.D., MINER T.L., NASH W.E., NELSON J.O., NHAN M.N., et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520–562.
- CHOPRA S., BRENDEN V., ZHANG J., AXTELL J.D. & PETERSON T. 1999. Molecular characterization of a mutable pigmentation phenotype and isolation of the first active transposable elements in *Sorghum bicolor*. *Proceedings of the National Academy of Sciences* 96: 15330–15335.
- CHUNG H., BOGWITZ M.R., MCCART C., ANDRIANOPOULOS A., FRENCH-CONSTANT R.H., BATTERHAM P. & DABORN P.J. 2007. Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. *Genetics* 175: 1071–1077.



- CLAEYS BOUUAERT C., LIPKOW K., ANDREWS S.S., LIU D. & CHALMERS R. 2013. The autoregulation of a eukaryotic DNA transposons. *eLife* 2: e00668.
- CLARK A.G., EISEN M.B., SMITH D.R., BERGMAN C.M., OLIVER B., MARKOW T.A., KAUFMAN T.C., KELLIS M., GELBART W., IYER V.N., et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450: 203–218.
- COATES B.S., HELLMICH R.L., GRANT D.M. & CRAIG A.A. 2012. Mobilizing the genome of Lepidoptera through novel sequence gains and end creation by non-autonomous Lep1 Helitrons. *DNA Research* 19: 11–21.
- CORDEAUX R. 2009. Gene conversion maintains nonfunctional transposable elements in an obligate mutualistic endosymbiont. *Molecular Biology and Evolution* 26: 1679–1682.
- CRESPI B. & NOSIL P. 2012. Conflictual speciation: species formation via genomic conflict. *Trends in Ecology and Evolution* 28: 48–57.
- DABOUSSI M.-J. & CAPY P. 2003. Transposable elements in filamentous fungi. *Annual Review of Microbiology* 57: 275–299.
- DE BOER J.G., YAZAWA R., DAVIDSON W.S. & KOOP B.F. 2007. Bursts of horizontal evolution of DNA transposons in the speciation of pseudotetraploid salmonids. *BMC Genomics* 8: 422.
- DELAURIÈRE L., CHÉNAIS B., HARDIVILLIER Y., GAUVRY L. & CASSE N. 2009. *Mariner* transposons as genetic tools in vertebrate cells. *Genetica* 137: 9–17.
- DERELLE E., FERRAZ C., ROMBAUTS S., ROUZE P., WORDEN A.Z., ROBBENS S., PARTENSKY F., DEGROEVE S., ECHEYNIÉ S., COOKE R., SAEYS Y., WUYTS J., JABBARI K., BOWLER C., PANAUD O., PIÉGU B., BALL S.G., RAL J.-P., BOUGET F.-Y., PIGANEAU G., DE BAETS B., PICARD A., DELSENY M., DEMAILLE J., VAN DE PEER Y. & MOREAU H. 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences* 103: 11647–11652.
- DIAO X., FREELING M. & LISCH D. 2006. Horizontal transfer of a plant transposon. *PLoS Biology* 4: 0119–0128.
- DOMBROSKI B., MATHIAS S., NANTHAKUMAR E., SCOTT A. & KAZAZIAN H. 1991. Isolation of an active human transposable element. *Science* 254: 1805–1808.
- FALKOWSKI P.G., BARBER R.T. & SMETACEK V. 1998. Biogeochemical controls and feedbacks on ocean primary production. *Science* 281: 200–206.
- FESCHOTTE C. & MOUCHÈS C. 2000. Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a *pogo*-like DNA transposon. *Molecular Biology and Evolution* 17: 730–737.
- FESCHOTTE C. & PRITHAM E.J. 2007. DNA transposons and the evolution of eukaryotic genomes. *Annual Review of Genetics* 41: 331–368.
- FILEE J., SIGUIER P. & CHANDLER M. 2007. Insertion sequence diversity in archaea. *Microbiology and Molecular Biology Reviews* 71: 121–157.
- FINNEGAN D.J. 1989. Eukaryotic transposable elements and genome evolution. *Trends in Genetics* 5: 103–107.
- FORTUNE P., ROULIN A. & PANAUD O. 2008. Horizontal transfer of transposable elements in plants. *Communicative and Integrative Biology* 1: 1–4.
- GILBERT C., SCHAACK S., PACE I.J.K., BRINDLEY P.J. & FESCHOTTE C. 2010. A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature* 464: 1347–1350.
- GIRAUD T. & CAPY P. 1996. Somatic activity of the *mariner* transposable element in natural populations of *Drosophila simulans*. *Proceeding of Royal Society of London B Biology and Sciences* 263: 1481–1486.
- GIRAUD T., FORTINI D., LEVIS C., LEROUX P. & BRYGOO Y. 1997. RFLP markers show genetic recombination in *Botryotinia fuckeliana* (*Botrytis cinerea*) and transposable elements reveal two sympatric species. *Molecular Biology and Evolution* 14: 1177–1185.
- GONZALEZ J., LENKOV K., LIPATOV M., MACPHERSON J.M. & PETROV D.A. 2008. High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biology* 6: 2109–2119.
- GONZALEZ J., KARASOV T.L., MESSER P.W. & PETROV D.A. 2010. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in *Drosophila*. *PLoS Genetic* 6: 1–13.
- GRANDBASTIEN M.A., LUCAS H., MOREL J.-B., MHIRI C., VERNHETTES S. & CASACUBERTA J. 1997. The expression of the tobacco Tnt1 retrotransposon is linked to plant defence responses. *Genetica* 100: 241–252.
- GRANUM E., RAVEN J.A. & LEEGOOD R.C. 2005. How do marine diatoms fix 10 billion tonnes of inorganic carbon per year? *Canadian Journal of Botany* 83: 898–908.
- GROVER C.E. & WENDEL J.F. 2010. Recent insights into mechanisms of genome size change in plants. *Journal of Botany* 2010: 1–8.
- HAAS B.J., KAMOUN S., ZODY M.C., JIANG R.H.Y., HANDSAKER R.E., CANO L.M., GRABHERR M., KODIRA C.D., RAFFAELE S., TORTO-ALABIDO T., BOZKURT O., AF-FONG A.M.V., ALVARADO L., ANDERSON V.L., MILES R. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461: 393–398.
- HARTL D. 2001. Discovery of the transposable element *mariner*. *Genetics* 157: 471–476.
- HASHIDA S.-N., KITAMURA K., MIKAMI T. & KISHIMA Y. 2003. Temperature shift coordinately changes the activity and the methylation state of transposon *Tam3* in *Antirrhinum majus*. *Plant Physiology* 132: 1207–1216.
- HASHIDA S.-N., UCHIYAMA T., MARTIN C., KISHIMA Y., SANO Y. & MIKAMI T. 2006. The temperature-dependent change in methylation of the *Antirrhinum* transposon *Tam3* is controlled by the activity of its transposase. *Plant Cell* 18: 104–118.
- HENSE W., BAINES J.F. & PARSCH J. 2007. X chromosome inactivation during *Drosophila* spermatogenesis. *PLoS Biology* 5: 2288–2295.
- HERMANN D. 2011. *Caractérisation d'éléments transposables de type mariner chez les microalgues marines*. PhD. Thesis, Maine University, Le Mans.
- HUA-VAN A., LE ROUZIC A., MAISONHAUTE C. & CAPY P. 2005. Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenetic and Genome Research* 110: 426–440.
- ILVES H., HÖRAK R. & KIVISAAR M. 2001. Involvement of  $\zeta$ S in starvation-induced transposition of *Pseudomonas putida* Transposon Tn4652. *Journal of Bacteriology* 183: 5445–5448.



- JAILLET J., GENTY M., CAMBEFORT J., ROUAULT J.D. & AUGÉ-GOUILLOU C. 2012. Regulation of *Mariner* transposition: The peculiar case of *Mos1*. *PLoS One* 7: 1–11.
- JIANG N., BAO Z., ZHANG X., EDDY2 S.R. & WESSLER S.R. 2004. Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431: 569–573.
- JOSEPH B., SCHLUETER J.A., CHANG DU J., GRAHAM M.A., MA J. & SHOEMAKER R.C. 2009. Retrotransposons within syntenic regions between soybean and *Medicago truncatula* and their contribution to local genome evolution. *The Plant Genome* 2: 211–223.
- KALENDAR R., TANSKANEN J., IMMONEN S., NEVO E. & SCHULMAN A.H. 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by *BARE-1* retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Sciences* 97: 6603–6607.
- KAPITONOV V.V. & JURKA J. 2004. Harbinger transposons and an ancient HARBII gene derived from a transposase. *DNA Cell Biology* 23: 311–324.
- KAPITONOV V.V. & JURKA J. 2005. RAG1 core and V(D)J recombination signal sequences were derived from *Transib* transposons. *PLoS Biology* 3: 311–324.
- KEMPEN F. & KÖCK U. 1998. Transposons in filamentous fungi: facts and perspectives. *Bioessays* 20: 652–659.
- KEMPEN F. & WINDHOFFER F. 2001. The hAT family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma* 110: 1–9.
- KHURANA J.S., XU J., WENG Z. & THEURKAUF W.E. 2010. Distinct functions for the *Drosophila* piRNA pathway in genome maintenance and telomere protection. *PLoS Genetics* 6: 1–8.
- KIDWELL M.G. 2002. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49–63.
- KIDWELL M.G. & LISCH D.R. 2001. Perspectives: transposable elements, parasitic DNA, and genome. *Evolution* 55: 1–24.
- KIMURA Y., TOSA Y., SHIMADA S., SOGO R., KUSABA M., SUNAGA T., BETSUYAKU S., ETO Y., NAKAYASHIKI H. & MAYAMA S. 2001a. OARE-1, a *Ty1-copia* retrotransposon in oat activated by abiotic and biotic stresses. *Plant and cell Physiology* 42: 1345–1354.
- KIMURA R.H., CHOUDARY P.V., STONE K.K. & SCHMID C.W. 2001b. Stress induction of *Bm1* RNA in silkworm larvae: SINES, an unusual class of stress genes. *Cell Stress Chaperone* 6: 263–272.
- KRAMEROV D.A. & VASSETZKY N.S. 2005. Short retrotransposons in eukaryotic genomes. *International Review of Cytology* 247: 165–221.
- LANDERS E.S., LINTON L.M., BIRREN B., NUSBAUM C., ZODY M.C., BALDWIN J., DEVON K., DEWARK D.M. & FITZHUGH W., et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.
- LECLERCQ S., GIRAUD I. & CORDAUX R. 2011. Remarkable abundance and evolution of mobile group II introns in *Wolbachia* bacterial endosymbionts. *Molecular Biology and Evolution* 28: 685–697.
- LE ROUZIC A. & CAPY P. 2005. The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. *Genetics* 169: 1033–1043.
- LEVY A.A. 2013. Transposons in plant speciation. In: *Plant transposons and genome dynamics in evolution* (Ed by N.V. Fedoroff), pp. 165–179. Wiley-Blackwell, Oxford.
- LOHE A.R., MORIYAMA E.N., LIDHOLM D.A. & HARTL D.L. 1995. Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. *Molecular Biology and Evolution* 12: 62–72.
- LOPEZ P., DESCLES J., ALLEN A.E. & BOWLER C. 2005. Prospects in diatom research. *Current Opinion in Biotechnology* 16: 180–186.
- LORENZI H., THIAGARAJAN M., HAAS B., WORTMAN J., HALL N. & CALER E. 2008. Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genomics* 9: 1–15.
- LYON, M.F. 2000. LINE-1 elements and X chromosome inactivation: a function for 'junk' DNA? *Proceedings of the National Academy of Sciences* 97: 6248–6249.
- MA L.J., VAN DER DOES H.C., BORKOVICH K.A., COLEMAN J.J., DABOUSSI M.J., DI PIETRO A., DUFRESNE M., FREITAG M., GRABHERR M., HENRISSAT B., et al. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464: 367–373.
- MACAS J. & NEUMANN P. 2007. Ogre elements—A distinct group of plant Ty3/Gypsy-like retrotransposons. *Gene* 390: 108–116.
- MACAS J., KOBLIZKOVA A. & NEUMANN P. 2005. Characterization of *Stowaway* MITEs in pea (*Pisum sativum* L.) and identification of their potential master elements. *Genome* 48: 831–839.
- MACHIDA C. & MACHIDA Y. 1989. Regulation of *IS1* transposition by the *insA* gene product. *Journal of Molecular Biology* 208: 567–574.
- MAHESWARI U., MOCK T., ARMBRUST V. & BOWLER C. 2009. Update of the Diatom EST Database: a new tool for digital transcriptomics. *Nucleic Acids Research* 37: D1001–D1005.
- MASLY J.P., JONES C.D., NOOR M.A.F., LOCKE J. & ORR H.A. 2006. Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* 313: 1448–1450.
- MAUMUS F., ALLEN A., MHIRI C., HU H., JABBARI K., VARDI A., GRANDBASTIEN M.-A. & BOWLER C. 2009. Potential impact of stress activated retrotransposons on genome evolution in a marine diatom. *BMC Genomics* 10: 1–19.
- MCCLEINTOCK B. 1951. Chromosome organization and genic expression. *Cold Spring Harbor Symposium on Quantitative Biology* 16: 13–47.
- MCCLEINTOCK B. 1984. The significance of responses of the genome to challenge. *Science* 226: 792–801.
- MERCHANT S.S., PROCHNIK S.E., VALLON O., HARRIS E.H., KARPOWICZ S.J., WITMAN G.B., TERRY A., SALAMOV A., FRITZ-LAYLIN L.K., MARÉCHAL-DROUARD L., MARSHALL W.F., QU L.-H., NELSON D.R., SANDERFOOT A.A., SPALDING M.H., KAPITONOV V.V., REN Q., FERRIS P., LINDQUIST E., SHAPIRO H., et al. 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318: 245–250.
- MEYER T.J., McLAIN AT, OLDENBURG J.M., FAULK C., BOURGEOIS M.G., CONLIN E.M., MOOTNICK A.R., DE JONG P.J., ROOS C., CARBONE L. & BATZER M.A. 2012. An Alu based phylogeny of gibbons (Hylobatidae). *Molecular Biology and Evolution* 29: 3441–3450.
- MHIRI C., MOREL J.B., VERNHETTES S., CASACUBERTA J.M., LUCAS H. & GRANDBASTIEN M.-A. 1997. The promoter of the tobacco *Tnt1* retrotransposon is induced by wounding and by abiotic stress. *Plant Molecular Biology* 33: 257–266.



- MONDEN Y., NAITO K., OKUMOTO Y., SAITO H., OKI N., TSUKUYAMA T., IDETA O., NAGAZAKI T., WESSLER S.R. & TANISAKA T. 2009. High potential of a transposon *mPing* as a marker system in *japonica* x *japonica* cross in rice. *DNA Research* 16: 131–140.
- MONTSANT A., JABBARI K., MAHESWARI U. & BOWLER C. 2005. Comparative genomics of the pennate diatom *Phaeodactylum tricornutum*. *Plant Physiology* 137: 500–513.
- MUÑOZ-DÍEZ C., VITTE C., ROSS-IBARRA J., GAUT B.S. & TENAILLON M.I. 2012. Using nextgen sequencing to investigate genome size variation and transposable element content. In: *Plant transposable elements*. Springer Berlin Heidelberg, 41–58 pp.
- NAGY Z. & CHANDLER M. 2004. Regulation of transposition in bacteria. *Research in Microbiology* 155: 387–398.
- NELSON D., TREGUER M., BRZEZINSKI P., LEYNAERT M.A. & QUEGUINER B. 1995. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Glob. Biogeochemistry Cycles* 9: 359–372.
- NORDEN-KRICHMAR T.M., ALLEN A.E., GAASTERLAND T. & HILDEBRAND M. 2011. Characterization of the small RNA transcriptome of the diatom, *Thalassiosira pseudonana*. *PLoS One* 6: 1–12.
- OGASAWARA H., OBATA H., HATA Y., TAKAHASHI S. & GOMI K. 2009. *Crawler*, a novel *Tc1/mariner*-type transposable element in *Aspergillus oryzae* transposes under stress conditions. *Fungal Genetics and Biology* 46: 441–449.
- OLIVER K.R. & GREENE W.K. 2009. Transposable elements: powerful facilitators of evolution. *Bioessays* 31: 703–714.
- OLIVER K.R. & GREENE W.K. 2012. Transposable elements and viruses as factors in adaptation and evolution: an expansion and strengthening of the TE-Thrust hypothesis. *Ecology and Evolution* 2: 2912–2933.
- OLIVER M.J., SCHOFIELD O. & BIDLE K. 2010. Density dependent expression of a diatom retrotransposon. *Marine Genomics* 3: 145–150.
- PALAZZOLI F., TESTU F.X., MERLY F. & BIGOT Y. 2010. Transposon tools: worldwide landscape of intellectual property and technological developments. *Genetica* 138: 285–299.
- PARISOD C., ALIX K., JUST J., PETIT M., SARILAR V., MHIRI C., AINOUCHE M., CHALHOUB B. & GRANDBASTIEN M.A. 2010. Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytologist* 186: 37–45.
- PARK M., PARK J., KIM S., KWON J.-K., PARK H.M., BAE I.H., YANG T.-J., LEE Y.H., KANG B.-C. & CHOI D. 2012. Evolution of the large genome in *Capsicum annuum* occurred through accumulation of single-type long terminal repeat retrotransposons and their derivatives. *The Plant Journal* 9: 1018–1029.
- PASYUKOVA E.G., NUZHIDIN S.V., MOROZOVA T.V., MACKAY T.F. 2004. Accumulation of transposable elements in the genome of *Drosophila melanogaster* is associated with a decrease in fitness. *Journal of Heredity* 95: 284–290.
- PATERSON A.H., BOWERS J.E., BRUGGMAN R., DUBCHAK I., GRIMWOOD J., GUNDLACH H., HABERER G., HELLSTEN U., MITROS T., POLIAKOV A., et al. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551–556.
- PEREIRA A., CUYPERS H., GIERL A., SCHWARZ-SOMMER Z., SAEDLER H. 1986. Molecular analysis of the En/Spm transposable element system of *Zea mays*. *EMBO Journal* 5: 835–841.
- PIEDNOËL M. & BONNIVARD E. 2009. DIRS1-like retrotransposons are widely distributed among Decapoda and are particularly present in hydrothermal vent organisms. *BMC Evolutionary Biology* 9: 1–15.
- PIEDNOËL M., ABERER A.J., SCHNEEWEISS G.M., MACAS J., NOVAK P., GUNDLACH H., TEMSH E.M. & RENNER S.S. 2012. Next-generation sequencing reveals the impact of repetitive DNA across phylogenetically closely related genomes of *Orobanchaceae*. *Molecular Biology and Evolution* 29: 3601–3611.
- PIÉGU B., GUYOT R., PICAULT N., ROULIN A., SANIYAL A., KIM H., COLLURA K., BRAR D.S., JACKSON S., WING R.A. & PANAUD O. 2006. Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Research* 16: 1262–1269.
- RAKOCEVIC A., MONDY S., TIRICHINE L., COSSON V., BROCARD L., IANTCHEVA A., CAYREL A., DEVIER B., EL-HEBA G.A.A. & RATET P. 2009. *MERE1*, a low-copy-number copia-type retroelement in *Medicago truncatula* active during tissue culture. *Plant Physiology* 151: 1250–1263.
- RAY D.A., FESCHOTTE C., PAGAN H.J.T., SMITH J.D., PRITHAM E.J., ARENSBURGER P., ATKINSON P.W. & CRAIG N.L. 2008. Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Research* 18: 632–639.
- ROBERTSON H.M. 1993. The *mariner* transposable element is widespread in insects. *Nature* 362: 241–245.
- ROBERTSON H.M. & LAMPE D.J. 1995. Recent horizontal transfer of a *mariner* transposable element among and between *Diptera* and *Neuroptera*. *Molecular Biology and Evolution* 12: 850–862.
- ROCAP G., LARIMER F.W., LAMERDLIN J., MALFATTI A., CHAIN P., AHLGREN A.A., COLEMAN M., HAUSER L., HESS W.R., JOHNSON Z.L., LAND M., LINDELL D., POST A.F., REGALA W., SHAH M., SHAW S.L., STEGLICH C., SULLIVAN M.B., TING C.S., TOLONEN A., WEBB E.A., ZINSER E.R. & CHISHOLM S.W. 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* 424: 1042–1047.
- ROSE M.R. & DOOLITTLE F. 1983. Molecular biological mechanisms of speciation. *Sciences* 220: 157–162.
- ROULIN A., PIÉGU B., WING R. & PANAUD O. 2008. Evidence of multiple horizontal transfers of the long terminal repeat retrotransposon *RIRE1* within the genus *Oryza*. *The Plant Journal* 53: 950–959.
- ROUND F.E., CRAWFORD R.M. & MANN D.G. 1990. *The diatoms: biology and morphology of the genera*. Cambridge University Press, Cambridge.
- RUBIN G.M., KIDWELL M.G. & BINGHAM P.M. 1982. The molecular basis of P-M hybrid dysgenesis: The nature of induced mutations. *Cell* 29: 987–994.
- SASAKURA Y., KONNO A., MIZUNO K., SATOH N., INABA K. 2008. Enhancer detection in the ascidian *Ciona intestinalis* with transposase expressing lines of *Minos*. *Developmental Dynamics* 237: 39–50.
- SCALA S., CARELS N., FALCIATORE A., CHIUSANO L.C. & BOWLER C. 2002. Genome properties of the diatom *Phaeodactylum tricornutum*. *Plant Physiol.* 129: 993–1002.



- SCHNABLE P.S., WARE D., FULTON R.S., STEIN J.C., WEI F., PASTERNAK S., LIANG C., ZHANG J., FULTON L., GRAVES T.A., MINX P., REILY A.D., COURTNEY L., KRUCHOWSKI S.S., TOMLINSON C., STRONG C., *et al.* 2009. The B73 Maize genome: complexity, diversity and dynamics. *Science* 326: 1112–1115.
- SHARMA A., SCHNEIDER K.L. & PRESTING G.G. 2008. Sustained retrotransposition is mediated by nucleotide deletions and interelement recombinations. *Proceedings of the National Academy of Sciences* 105: 15470–15474.
- SHIRASAWA K., HIRAKAWA H., TABATA S., HASEGAWA M., KIYOSHIMA H., SUZUKI S., SASAMOTO S., WATANABE A., FUJISHIRO T. & ISOBE S. 2012. Characterization of active miniature inverted-repeat transposable element in the peanut genome. *Theoretical and Applied Genetics* 124: 1429–1438.
- SILVA J.C. & KIDWELL M.G. 2004. Evolution of *P* elements in natural populations of *Drosophila willistoni* and *D. sturtevantii*. *Genetics* 168: 1323–1335.
- SINZELLE L., IZSVÁK Z. & IVICS Z. 2009. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cellular and Molecular Life Sciences* 66: 1073–1093.
- SLOTKIN R.K. & MARTIENSSSEN R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics* 8: 272–285.
- SMETACEK V. 1999. Diatom and the ocean Carbonic cycle. *Protist* 150: 25–32.
- SMIT A.F. 1996. The origin of interspersed repeats in the human genome. *Current Opinion in Genetics and Development* 6: 743–748.
- STRAND D.J. & MC DONALD J.F. 1985. *Copia* is transcriptionally responsive to environmental stress. *Nucleic Acids Research* 13: 4401–4410.
- SUH A., KRIEKS J.O., DONNELLAN S., BROSIUS J. & SCHMITZ J. 2012. A universal method for the study of CR1 Retrotransposons in nonmodel bird genomes. *Molecular Biology and Evolution* 29: 2899–2903.
- TAKAHASHI R., MORITA Y., NAKAYAMA M., KANAZAWA A. & ABE J. 2012. An Active CACTA-Family transposable element is responsible for flower variegation in wild soybean *Glycine soja*. *The Plant Genome*, 5: 62–70.
- TAKEDA S., SUGIMOTO K. & HIROCHIKA H. 1999. A 13-bp cis-regulatory element in the LTR promoter of the tobacco retrotransposon *Tio1* is involved in responsiveness to tissues culture, wounding, methyl jasmonate and fungal elicitor. *The Plant Journal* 18: 383–393.
- TAPIA G., VERDUGO I., YANEZ M., AHUMADA I., THOEDULOZ C., CORDERO C., POBLETE F., GONZALEZ E. & RUIZ-LARA S. 2005. Involvement of ethylene in stress-induced expression of the *TLC11* retrotransposon from *Lycopersicon chilense* Dun. *Plant Physiology* 138: 1596–1599.
- TESTORI A., CAIZZI L., CUTRUPI S., FRIARD O., DE BORTOLI M. & CASELLE M. 2012. The role of transposable elements in shaping the combinatorial interaction of transcription factors. *BMC Genomics* 13: 1–16.
- TIAN Z., YU Y., LIN F., YU Y., SANMIGUEL P.J., WING R.A., MCCOUCH S.R., MA J.M. & JACKSON S.A. 2011. Exceptional lability of a genomic complex in rice and its close relatives revealed by interspecific and intraspecific comparison and population analysis. *BMC Genomics* 12: 1–12.
- TIRICHINE L. & BOWLER C. 2011. Decoding algal genomes: tracing back the history of photosynthetic life on Earth. *Plant Journal* 66: 45–57.
- TREANGEN T.J., ABRAHAM A.L., TOUCHON M. & ROCHA E.P. 2009. Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS Microbiology Reviews* 33: 539–571.
- VITTE C. & PANAUD O. 2005. Retrotransposon and flowering plant genomes size: emergence of a increase/decrease model. *Cytogenetic and Genome Research*, 110: 91–107.
- VITTE C., PANAUD O. & QUESNEVILLE H. 2007. LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss. *BMC Genomics* 8: 1–15.
- WALLAU G.L., HUA-VAN A., CAPY P., LORETO E.L.S. 2011. The evolutionary history of *mariner*-like elements in Neotropical drosophilids. *Genetica* 139: 327–338.
- WANG J., STATEN R., MILLER T.A. & PARK Y. 2005. Inactivated *mariner*-like elements (*MLE*) in pink bollworm, *Pectinophora gossypiella*. *Insect Molecular Biology* 14: 547–553.
- WARD P.N., HOLDEN M.T., LEIGH J.A., LENNARD N., BIGNELL A., BARRON A., CLARK L., QUAIL M.A., WOODWARD J., BARELL B.G., EGAN S.A., FIELD T.R., MASKELL D., KELHOE M., DOWSON C.G., CHANTER N., WHATMORE A.M., BENTLEY S. & PARKNILL J. 2009. Evidence for niche adaptation in the genome of the bovine pathogen *Streptococcus uberis*. *BMC Genomics* 10: 1–17.
- WESSLER S.R., BUREAU T.E. & WHITE S.E. 1995. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Current Opinion in Genetics and Development* 5: 814–821.
- WICKER T., SABOT F., HUA-VAN A., BENNETZEN J., CAPY P., CHALOU B., FLAVELL A., LEROY P., MORGANTE M., PANAUD O., PAUX E., SANMIGUEL P. & SCHULMAN A. 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* 8: 973–982.
- WRIGHT S. & FINNEGAN D. 2001. Genome evolution: sex and the transposable element. *Current Biology* 11: R296–R298.
- YANDELL M. & ENCE D. 2012. A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics* 13: 329–342.
- YEADON P.J., CATCHESIDE D.B.A. 1995. *Guest*: a 98 bp inverted repeat elements in *Neurospora crassa*. *Molecular Genomics and Genetics* 247: 105–109.
- ZERBIB D., PRENTKI P., GAMAS P., FREUND E., GALAS D.J. & CHANDLER M. 1990. Functional organization of the ends of IS1: specific binding site for an IS1-encoded protein. *Molecular Microbiology*, 4: 1477–1486.
- ZHANG L., DAWSON A. & FINNEGAN D.J. 2001. DNA-binding activity and subunit interaction of the *mariner* transposase. *Nucleic Acids Research* 29: 3566–3575.
- ZUCCOLO A., SEBASTIAN A., TALAG J., YU Y., KIM H., COLLURA K., KUDRNA D. & WING R.A. 2007. Transposable element distribution, abundance and role in genome size variation in the genus *Oryza*. *BMC Evolutionary Biology* 7: 152–167.

ORIGINAL PAPER

# First Evidence of Mariner-like Transposons in the Genome of the Marine Microalga *Amphora acutiuscula* (Bacillariophyta)



Duc Hung Nguyen<sup>a</sup>, Dorothée Hermann<sup>a</sup>, Aurore Caruso<sup>a</sup>, Emmanuelle Tastard<sup>a</sup>, Justine Marchand<sup>a</sup>, Jacques-Déric Rouault<sup>b</sup>, Françoise Denis<sup>a,c</sup>, Stanislas Thiriet-Rupert<sup>a</sup>, Nathalie Casse<sup>a,1,2</sup>, and Annick Morant-Manceau<sup>a,1,2</sup>

<sup>a</sup>PRES LUNAM université, Université du Maine, Laboratoire Mer, Molécules, Santé (EA 2160), avenue Olivier Messiaen, 72085 Le Mans cedex 9, France

<sup>b</sup>Laboratoire Evolution Génome et Spéciation, CNRS UPR 9034, avenue de la terrasse, bâtiment 13, BP 1, 91198 Gif-sur-Yvette, France

<sup>c</sup>UMR BOREA « Biologie des Organismes et Ecosystèmes Aquatiques », Muséum National d'Histoire Naturelle (MNHN), CP 26, 43 rue Cuvier, 75231 Paris Cedex 05, France

Submitted April 25, 2014; Accepted August 18, 2014  
Monitoring Editor: Michael Melkonian

**Mariner-like elements (MLEs)** are transposable elements able to move in the host genomes by a “cut and paste” mechanism. They have been found in numerous organisms. We succeeded in amplifying complete and truncated MLEs in the marine diatom *Amphora acutiuscula*. Full-length MLEs of 2,100 bp delimited by Imperfect Terminal Inverted Repeats revealed an Intact Open Reading Frame, suggesting that the MLEs could be active. The DNA binding domain of the corresponding putative transposase could have two Helix-Turn-Helix and a Nuclear Location Site motifs, and its catalytic domain includes a particular triad of aspartic acids DD43D not previously reported. The number of copies was estimated to be 38, including approximately 20 full-length elements. Phylogenetic analysis shows that these peculiar MLEs differ from plant and other stramenopile MLEs and that they could constitute a new sub-family of *Tc1-mariner* elements.

© 2014 Elsevier GmbH. All rights reserved.

**Key words:** Diatom; *mariner*-like element; DNA transposon; transposable elements.

## Introduction

Transposable elements (TEs) are repeated mobile DNA sequences found in almost all genomes (Carr et al. 2008; Kidwell and Lisch 2001; Robertson 2002). Within the chromatin, TEs move from one locus to another by a process known as transposition. Genomes can be invaded by TEs to

<sup>1</sup>Corresponding authors; fax +33 243833917

<sup>2</sup>These authors contributed equally to this work.  
e-mail [ncasse@univ-lemans.fr](mailto:ncasse@univ-lemans.fr) (N. Casse),  
[annick.morant-manceau@univ-lemans.fr](mailto:annick.morant-manceau@univ-lemans.fr) (A. Morant-Manceau).

<http://dx.doi.org/10.1016/j.protis.2014.08.003>  
1434-4610/© 2014 Elsevier GmbH. All rights reserved.



varying extents, meaning that the proportion of the DNA that they occupy is very variable, ranging for instance from 2% in bacteria or yeast to over 90% in some plants (Bennetzen 2005; de Koning et al. 2011; Kidwell 2002; Lander et al. 2001; Lönning and Saedler 2002; Schnable et al. 2009; Zuccolo et al. 2007). Most of the TEs in a genome are inactive, either as a result of the accumulation of mutations, or of being silenced by epigenetic mechanisms (Slotkin and Martienssen 2007). However, some of them are active and might transpose under stressful conditions (Capy et al. 2000; Grandbastien et al. 2005; Maumus et al. 2009). If they do transpose they could induce gene inactivation by being inserted directly into a gene or its promoter. Most of the effects of TEs are deleterious, but their movement can also generate genetic diversity which permits adaptation to environmental change (Biémont and Viera 2006; Capy 2004; Kidwell and Lisch 1997; Laudencia-Chingcuanco and Fowler 2012; Lönning and Saedler 1997; Silva et al. 2005). TEs generate genomic plasticity and so are viewed as key players in genome evolution (Böhne et al. 2008; Kidwell and Lisch 1997; Tenaillon et al. 2010; Venner et al. 2009). Because of their ability to cut, move and paste DNA fragments, some TEs have been studied in relation to the development of genetic tools for mutagenesis and transgenesis (Finnegan 1992; Ivics et al. 2009).

Two classes of TEs have been distinguished on the basis of their structural organization and their transposition mechanism (Finnegan 1989, 1992; Kapitonov and Jurka 2008; Wicker et al. 2007). Class I elements, or retrotransposons, transpose via RNA intermediates that are reverse-transcribed and integrated as supplementary copies into the host genome. This type of transposition is very invasive as the number of retrotransposon copies increases over time. Class II elements, also known as DNA transposons, transpose mainly, but not exclusively, via a "cut-and-paste" mechanism. Wicker et al. (2007) have divided class II into two subclasses: subclass I consists of elements with Terminal Inverted Repeats (TIRs) and Crypton elements that have neither TIRs, nor a transposase gene (Goodwin et al. 2003); subclass II includes Helitron elements which use a rolling circle process to move (Kapitonov and Jurka 2001) and Maverick elements, also called *Polintons*, which are large TEs (Feschotte and Pritham 2005; Kapitonov and Jurka 2006). TIR elements contain at least an Open Reading Frame (ORF) that encodes a transposase. The ORF is flanked by untranslated sequences (UTRs) and TIRs. These latter are recognized by the transposase, which excises the transposon at

its extremities and inserts it at a new target site. The target site is duplicated at the end of the transposition process, which is known as Target Site Duplication (TSD). The TSD is also used as a diagnostic feature to characterize the DNA transposons (Kapitonov and Jurka 2008; Wicker et al. 2007). Plant and mammalian genomes are dominated by retrotransposons (Deininger and Roy-Engel 2002; Kentner et al. 2003; Vitte et al. 2007). However, class II elements, which are less invasive than retrotransposons, constitute the majority of TEs in the genomes of the bat *Myotis lucifugus* and in some fungi (Daboussi and Capy 2003; Feschotte and Pritham 2007; Hua-Van et al. 2000; Pace et al. 2008; Ray et al. 2008).

*Tc1-mariner* is one of the most widespread super-families of TIR elements within eukaryotic genomes (Plasterk et al. 1999; Wicker et al. 2007). *Tc1*-like elements (*TLEs*) and *mariner*-like elements (*MLEs*) display common motifs with the *Tc1* element in *Caenorhabditis elegans* (Henikoff and Plasterk 1988) and *mariner* in *Drosophila mauritiana* (Jacobson et al. 1986), respectively. *TLEs* and *MLEs* families have been associated with bacterial Insertion Sequences *IS630* by Shao and Tu (2001), hence the name of the superfamily *ITm* (*IS630-Tc1-mariner*). Wicker et al. (2007) also designated the *Tc1-mariner* superfamily as *DTT* (DNA transposon, TIR group, *Tc1-mariner*). This TIR superfamily is characterized by a dinucleotide TA as TSD. The transposases of *Tc1-mariner* elements have an N-terminal domain containing one or two helix-turn-helix (HTH) motifs that recognize and bind the TIRs, and a C-terminal domain containing a conserved catalytic triad (DDxE for *TLEs* and DDxD for *MLEs*) (Brillet et al. 2007; Doak et al. 1994; Plasterk et al. 1999). In different studies, the number of amino residues (34, 37, 38 39, and 41) located between the last two amino acids of the triad is a characteristic used to classify these elements within families called ITm DDxD/E (Brillet et al. 2007; Coy and Tu 2005; Shao and Tu 2001). However, this criterion is not used in some other classifications such as that of Wicker et al. (2007) because in many groups the catalytic domains are not well characterized. Wicker et al. (2007) propose the DNA sequence similarity as the criterion to identify the families.

In the ocean, diatoms (Bacillariophyta) make up 40% of the phytoplanktonic eukaryote species, and are responsible for around 20% of the total worldwide primary production (Field et al. 1998; Granum et al. 2005; Simon et al. 2009). Diatoms belong to the Stramenopile (or Heterokonta) taxon (Baldauf 2008; Burki et al. 2008; Cavalier-Smith 1999; Hackett et al. 2007), which includes the



photosynthetic Bacillariophyta and also parasites such as the oomycete *Phytophthora*, which is responsible for “potato blight”. Red algae, Chlorophyta, and land plants probably all originate from a primary endosymbiosis event during the Proterozoan era; a cyanobacterium engulfed by a heterotrophic cell could have evolved into a chloroplast after gene transfer between the prokaryote and the host cell (Falkowski et al. 2004; Keeling et al. 2004; Reyes-Prieto et al. 2007). Then, at the end of the Precambrian era, a red alga may have been engulfed (secondary endosymbiosis) by a eukaryotic cell and evolved as a chloroplast in the ancestor of the Chromalveolata, this latter group including the Stramenopiles (Keeling et al. 2004; Nisbet et al. 2004; Parfrey et al. 2006; Seidl et al. 2012; Simon et al. 2009).

Genome sequencing of four marine diatoms, *Thalassiosira pseudonana* (Armbrust et al. 2004), *Phaeodactylum tricornutum* (Bowler et al. 2008), *Fragilariopsis cylindrus* (JGI databank), and *Pseudo-nitzschia multiseries* (JGI databank) has confirmed that the nuclear genome of diatoms is a mosaic of genes coming from two nuclei (that of the heterotrophic host and that of the red alga) and from a cyanobacterium (Armbrust et al. 2004; Bowler et al. 2008; Deschamps and Moreira 2012; Lane and Archibald 2008; Tirichine and Bowler 2011). TEs represent a genomic fraction of 1.1 and 5.8% in *T. pseudonana* and *P. tricornutum* respectively, which is low compared to maize (80%) or human (45%) TEs. These TEs are essentially *Ty1-copia* retrotransposons, which account for 90% of the TEs in *P. tricornutum* (Bowler et al. 2008; Maumus et al. 2009). Recently, a TIR transposon, *Costa*, has been detected in silico in the diatom *P. multiseries* (Maurus et al. unpublished data). Hermann (2011) searched for MLEs in various marine diatom species, and revealed the presence of several partial sequences of these transposons in the genome of *Amphora acutiuscula*. The aim of the present study was to characterize complete MLE transposons in the genome of *A. acutiuscula*.

## Results

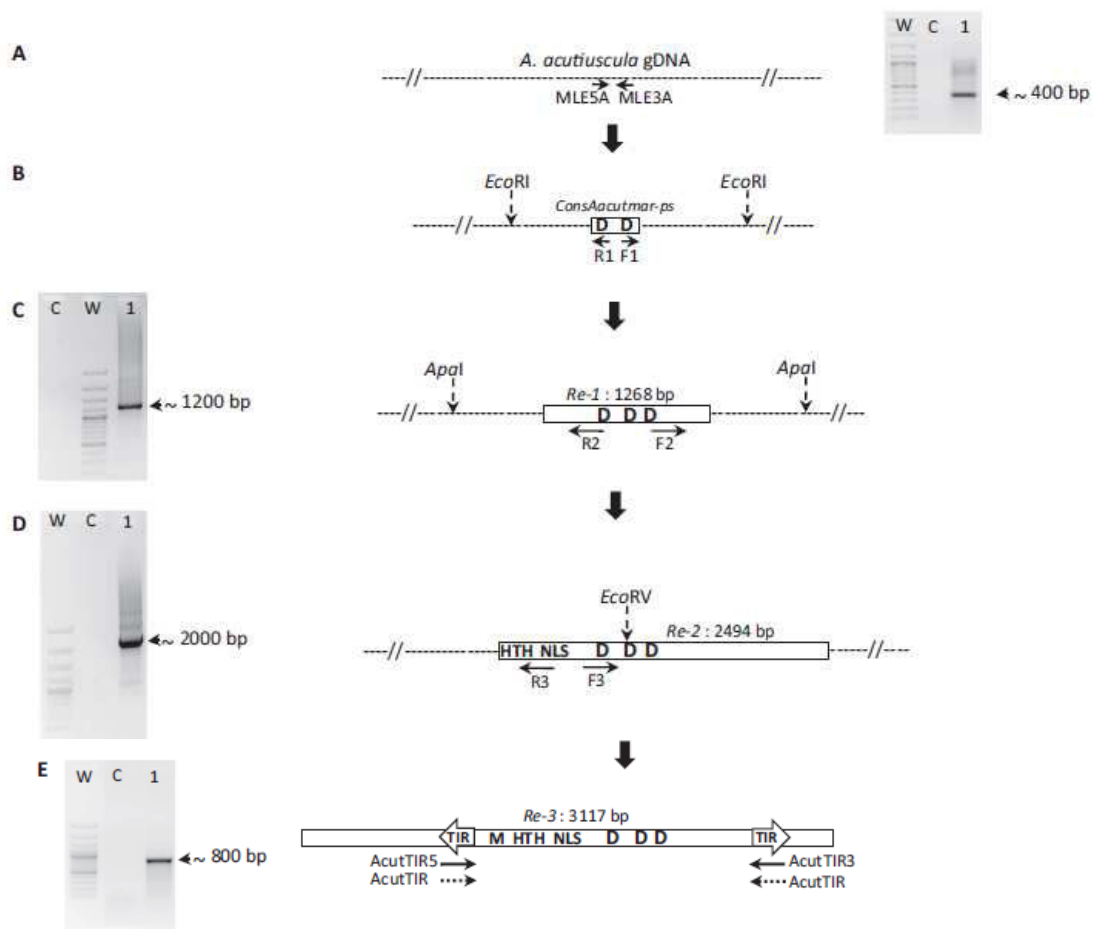
### Detection of MLEs in the Genome of *A. acutiuscula* by PCR

Since one MLE has been found in silico in one of the four sequenced genomes of diatoms (*Costa*, *P. multiseries*), we designed a primer on its TIRs and used it in order to amplify potential full-length elements of

*A. acutiuscula* by PCR. Since these attempts failed, we then used the degenerate primers MLE5A and MLE3A that allowed to detect MLEs in plants (Feschotte and Wessler 2002). The amplification by PCR produced a single fragment approximately 380 bp in length in the genome of *A. acutiuscula* (Fig. 1A). This fragment was cloned and then, eight clones were sequenced. After analysis by BlastX, each clone showed an identity close to 50% with the corresponding part of the MLE transposase protein sequence of plants such as *Avena sativa*, *Pisum sativum*, or *Phyllostachys edulis*. The eight sequences detected in *A. acutiuscula* were called *Aacutmar-ps 1.1* to *1.8* (*A. acutiuscula mariner partial sequence 1. clones 1 to 8*, accession numbers: KJ751531 to KJ751538). They were 372 to 379 bp in length, seven of them exhibited neither a stop codon nor a frameshift. Pairwise clone comparison indicated sequence identities ranging from 80 to 99% at the nucleotide level, and similarities from 80 to 100% at the amino acid level. The eight clones were used to create a unique consensus designated *ConsAacutmar-ps*, which was used as the first template for inverse PCR (iPCR).

### Full-length MLE Amplification Using Inverse PCR

To obtain a full-length MLE, three rounds of iPCR were necessary, each using the sequence elongated in the previous run as a template to design the primer sets and a restriction enzyme for further elongation. The first round of iPCR was carried out using *ConsAacutmar-ps*, enzyme *EcoRI* for cutting outside the sequence, and the nested primers F1 (in and out) and R1 (in and out; Table 1, Fig. 1B). This first round of iPCR amplified an approximately 1200-bp long fragment that was in silico assembled with *ConsAacutmar-ps* to yield a reconstructed element of 1,268 bp, designated *Re-1* for Reconstructed sequence 1 (Fig. 1C). The analysis of *Re-1* by blastX showed a similarity close to 45% with MLE transposases of *Beta vulgaris* and the marine diatom *Pseudo-nitzschia multiseries*. In silico amino acid translation of *Re-1* exhibited the three motifs of the catalytic triad of the MLE transposase. The first D was included in a motif “VDEKWF” slightly modified from the canonical “IDEKWF”, the second D in “IQQDNA”, and the third D in “SPDTNINDLAFF”. The last two motifs are identical to the canonical ones. The distance between the second and the third D is 43 amino acids (DD43D), which differs from the classical DD39D of MLEs in plants. At the 5' end, no HTH or NLS could be detected indicating the need for



**Figure 1.** A) PCR amplification of partial sequences of *MLE* using the degenerate primers MLE5A and MLE3A and the resulting amplified fragments visualized on a 1% agarose gel. B) Diagram of *ConsAactumar-ps*, the consensus partial sequence resulting after PCR A and indicating the cut-off position of the enzyme (*EcoRI*) and the location of the specific primers (R1/F1) for the first run of iPCR. C) Fragment amplified after iPCR B and diagram of the *in silico* reconstructed fragment Re-1 indicating the cut-off position of the enzyme (*Apal*) and the location of the primers (R2/F2) for the second run of iPCR. D) Fragment amplified after iPCR C and diagram of the *in silico* reconstructed fragment Re-2 indicating the cut-off position of the enzyme (*EcoRV*) and the location of the primers (R3/F3) for the third run of iPCR. E) Fragment amplified after iPCR D and diagram of the *in silico* reconstructed sequence Re-3 indicating also the location of the primers designed on TIRs to amplify the full-length elements by PCR. W: molecular weight marker, 100 pb DNA ladder plus (Promega); C: control PCR without gDNA. TIR: Terminal Inverted Repeat. Small black arrows and dashed arrows: locations of the specific and degenerate primers. D (in bold): aspartic amino acid of the catalytic core of the putative transposase; HTH: Helix Turn Helix; M: start codon; NLS: nuclear localization signal.

further extension in this region. The sequence *Re-1* was used as a template for further extension by a second round of iPCR using the enzyme *Apal* (which does not cut within *Re-1*) and the nested primers R2 (in and out) and F2 (in and out; Table 1).

The iPCR amplification gave a single 2000-bp long band that was assembled *in silico* with *Re-1* and yielded a fragment of 2,494 bp designated *Re-2* (Fig. 1D). After translation, two putative HTH and one putative NLS were found to be present, but still

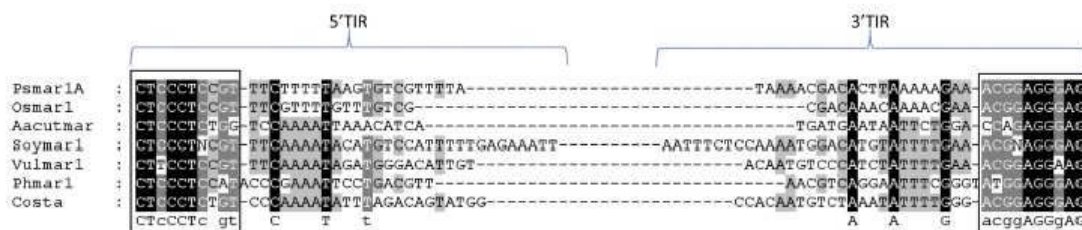


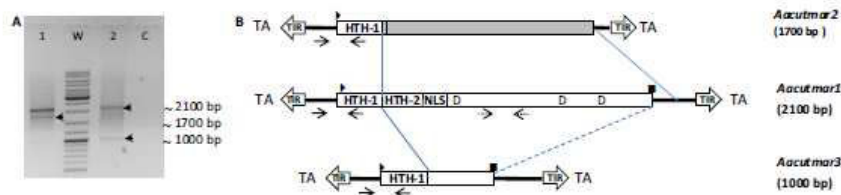
**Table 1.** The primers used for PCR, iPCR and copy number estimation, and the corresponding annealing temperature.

PCR primers	Annealing temperature (°C)
MLE5A: ATHGATGARAARTGGTTC	47
MLE3A: GCATTRTCYTGTDAT	
Acut-TIR5: TACTCCCTCTGGTCCAAAATTA	58
Acut-TIR3: TACCCCTCTGGTCCAGAATTA	
Acut-TIR: TACYCCCTCTGGTCCARAATTA	58
Costa TIR: TACTCCCTCYGTCCCAAAT	56-58
<b>Inverse PCR primers</b>	
R1out: GCACACAAGAAC (A/G) TCAC(T/C) TTGC	55
F1out: TTGG(C/T) AT(T/A) TGGCCATTGGGA(C/T) ATG	
R1int: CAAGAAC (A/G) TCAC(T/C) TTGCCAATGTA	55
F1int: TGGCCATTGGATATGTTGG	
R2out: AAGTACTTCGGACTTCCATTCC	55
F2out: TCAATCACCAGACACAAACATC	
R2int: ACTTCCATTCTTACCATATTTGGG	57
F2int: ACACAAACATCAATGATTTGGCATTC	
R3out: TCTCCTACACAGCGCGGTT	53
F3out: GTGGATGAAGATGAGATCAAATC	
R3int: AGCGCGGTTGTGCCTCAATT	55
F3int: GATCAAATCAAAGAAACATGAGAGAG	
<b>Primers for copy number estimation</b>	
AcutmarF1: AAAGTGATGTTCTTGTGTG	60
AcutmarR1: CCAACATATCCAAATGGCCAAAT	
AcutmarF2: GCTGGAACCTCGCACACAAT	60
AcutmarF2: ACAGCTCTCCYACACAGCG	

no start codon. A third iPCR round was performed to carry on elongating the 5' region of the transposon with *Re-2* as a template, EcoRV (that cleaves within the transposase gene), and primers R3 and F3 that are located near the 5' region (Fig. 1D). The resulting sequence was aligned with the 5' end of *Re-2* and it extended the 5' region by 620 bp. The in silico assembly of both sequences gave rise to a reconstructed sequence 3,117 bp in length, designated *Re-3* (Fig. 1E). In silico translation of *Re-3* made it possible to detect a start codon. In the region located outside the transposase gene,

the TIRs were investigated using the appropriate algorithm (see Methods: Analysis of diatom *MLE* sequences). After analyses, the 5' and 3' TIRs were detected. According to the sequence *Re-3*, the conceptual full-length *MLE* transposon exhibited a full size of 2126 bp, two imperfect 27-bp TIRs containing four differences at positions 2, 15, 21, and 22, respectively. This element is flanked at each extremity by the characteristic *MLE*-TSD: a dinucleotide TA. The nucleic comparison of the *Acutmar* TIRs with those of six full-length *MLE*s of plants and of a diatom collected from the databank

**Figure 2.** Alignment of 5' and 3' TIR sequences of *Acutmar1* and *MLE*s from databanks. The conservative sequence at the end of each TIR is boxed. The dashed lines indicate the remaining *MLE* sequence. Costa JX013907; Osmar1 AF172282; Phmar1 HM581665; Psmar1 AY8333549; Soymar1 AF078934; Vulmar1 AJ556159.



**Figure 3.** A) Full-length sequences of MLEs amplified by PCR on a 1% agarose gel. Lane 1: PCR with the specific primers AcutTIR5 and AcutTIR3; Lane 2: PCR with the degenerate primer AcutTIR; W: molecular weight marker, 100-pb DNA ladder plus (Promega); C: control PCR without gDNA. B) Schematic representations of the MLE sequences amplified on A). TA: Target Site Duplication; TIR: Terminal Inverted Repeat; black triangle: start codon; HTH 1 and 2: Helix Turn Helix; NLS: Nuclear Localization Signal; D: aspartic acid of the catalytic core of the putative transposase; black square: stop codon; black arrows and dashed arrows: PCR primers used respectively for *Aacutmar* and *Aacutmar1* copy number estimation. In *Aacutmar2*, the gray box indicates that this part of the element contains numerous scattered deletions.

(*Soymar1* from soybean *Glycine max*, *Osmar1* from the rice *Oryza sativa*, *Vulmar1* from the beet *Beta vulgaris*, *Psmar1* from the pea *Pisum sativum*, *Phmar1* from the bamboo *Phyllostachys edulis*, *Costa* from the diatom *P. multiseriata*) revealed that the first ten nucleotides of TIRs are relatively conserved between these species, and gave a slightly degenerated sequence of 5'-CTcCCTcYgt-3' (Fig. 2). Since the Re-3 sequence was an in silico assemblage, we then verified that it was indeed present in the genome of *A. acutiuscula*.

#### Confirmation of the Presence of Full-length MLEs in the Host Genome

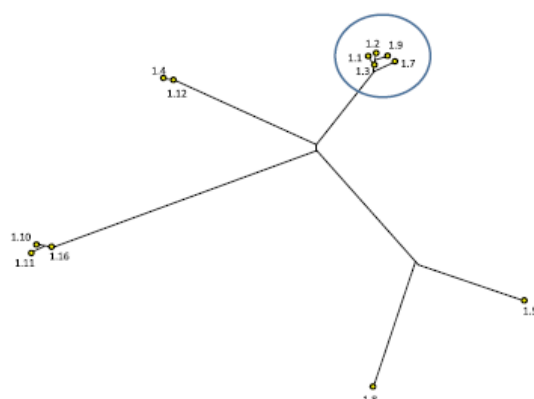
In order to amplify full-length MLEs, the TIRs of Re-3 were used to design specific and degenerate primers (Fig. 1E; Table 1). After PCR amplification, two bands were detected. The PCR products were found to be approximately 2,100 and 1,700 bp long when using specific primers, and 2,100 and 1,000 bp long using the degenerate primers (Fig. 3A). Each band was cloned and several clones were sequenced. The blast analyses revealed that the fragments 2,100 bp in length were full-length MLE sequences, and the 1,700- and 1,000-bp fragments were elements containing deletions. The complete element was designated *Aacutmar1* (for *A. acutiuscula mariner-like transposon 1*, accession numbers: KJ751515 to KJ751526), and the 1,700 and 1,000 bp deleted sequences were designated *Aacutmar2* and *Aacutmar3*, respectively (Fig. 3). In *Aacutmar2* (accession numbers: KJ751527 and KJ751528), numerous deletions, more than 40 ranging from 3 to 50 bp, were scattered throughout the length of the transposase gene and the 3' end of the element. *Aacutmar3* (accession numbers: KJ751529 and KJ751530) is

a deleted element that has lost more than 1000 bp of the transposase gene as a result of four large, 120- to 360-bp deletions (Fig. 3B).

#### Polymorphism of *Aacutmar1*

To estimate the polymorphism of *Aacutmar1* transposon in *A. acutiuscula*, twelve sequences were isolated from four independent PCRs (three *Aacutmar1* clones were collected after each of the four amplifications). The resulting sequences were designated *Aacutmar1.1* to *1.12*, they ranged in length from 2,094 to 2,132 bp. Pairwise comparison of the nucleotide sequences of the 12 clones indicated identities of more than 90% of the transposase gene sequences, and similarities of more than 80% between the putative transposases. The conceptual translation also revealed that six of these elements possessed an open reading frame (ORF), indicating that they could encode a functional transposase. The polymorphism of the twelve *Aacutmar1* clones was studied using the Fluxus program, allowing us to draw a network representing the relationships between the different nucleic sequences (Fig. 4). The *Aacutmar 1.1*, *1.2*, *1.3*, *1.7*, and *1.9* sequences, which contained an intact ORF clustered together, and displayed 99% similarity by pairwise comparison. *Aacutmar1.12* and *Aacutmar1.4* formed a second group. *Aacutmar1.12* exhibited two short insertions and one deletion conserving a potential ORF, but it displayed only 90-94% similarity with sequences of the first cluster at the nucleotide level. However, although *Aacutmar1.4* possesses an intact reading frame spanning the first 439 amino acids, it has lost the last 50 residues of the translated sequence. The five other clones (*1.5*, *1.6*, *1.8*, *1.10*, and *1.11*) have accumulated mutations leading to stop codons and





**Figure 4.** Haplotype network showing the relationships between the 12 haplotypes *Aacutmar1* isolated in *A. acutiuscula*. Each circle corresponds to one sequence.

frameshifts in the transposase gene and throughout the sequence.

### Structural Characteristics of *Aacutmar1* Transposons

The twelve sequences were used to create a 2,114-bp consensus designated *ConsAacutmar1*, with 27-bp TIRs, a 5' UTR of 152bp and a 3' UTR of 435bp. *ConsAacutmar1* possesses a 1,473-bp long ORF that potentially might encode a 491 amino acid transposase, named *AaTpase* (for *A. acutiuscula* transposase). The putative *AaTpase* was compared to similar proteins from data-banks. The most closely-related sequences are *Psmar1* (*P. sativum*) with 34% of similarity, *Costa* (*P. multiseri*) (33%), *Soymar1* (*G. max*) (31%), *Osmar1* (*O. sativa*) (30%), *Phmar1* (*P. edulis*) (28%), and *Vulmar1* (*B. vulgaris*) (27%). The protein sequences of the transposase of *Aacutmar1* and those in the data banks were aligned to detect any characteristic motifs in these enzymes (Fig. 5). *AaTpase* is an enzyme with a DNA-binding domain and a catalytic core. The first domain contains two putative helix turn helix motifs (HTH) at positions 59-79 and 129-149, a highly-conserved three amino acid motif, GRK (position 105-107), and a putative nuclear localization signal (NLS) at position 45 to 60. The detection of a NLS indicates that *AaTpase* is putatively a nuclear enzyme. As mentioned above, the catalytic domain exhibits a catalytic core, DD43D, in which each D is included in the motifs DEKWF, IQQD, and SPDTNINDLAFF, respectively, which are slightly modified from the highly-conserved motifs detected in plant MLEs.

### Phylogenetic Relationships of *Aacutmar1* among MLEs

The phylogenetic relationships between *Aacutmar1* and the other DNA transposons belonging to the *Tc1-mariner* superfamily have been investigated. In the resulting unrooted tree, the characteristic of the catalytic triad (DDxD) has been mentioned (Fig. 6). The phylogenetic tree exhibits two main clades. One of them includes the large monophyletic DD34D group, which contains MLEs isolated from animals (insects, crabs, *Caenorhabditis*), human or the protist *Trichomonas vaginalis*. The other one groups together MLEs isolated from the green alga *Ostreococcus tauri*, plants (DD39D), and Stramenopiles (the oomycete *Phytophthora infestans* and the diatoms *P. multiseri* and *A. acutiuscula*). Within this clade, the MLEs isolated from *O. tauri* form a separate cluster. The other ones are subdivided in two clusters: one gathers together MLEs from plants and the other one gathers together those in Stramenopiles (*A. acutiuscula*, *P. multiseri* and *P. infestans*). Within this cluster, the MLEs from *A. acutiuscula* could form a group apart from *P. multiseri* and *P. infestans*.

### Estimation of the Number of *Aacutmar* Copies

The number of *Aacutmar* copies was estimated on the basis of the results obtained from four different gDNA samples. The number of full-length *Aacutmar1* copies in the *A. acutiuscula* genome was estimated to be  $21 \pm 8$ , and the total number of the three *Aacutmar* (1, 2, and 3) elements was estimated to be  $38 \pm 14$ .

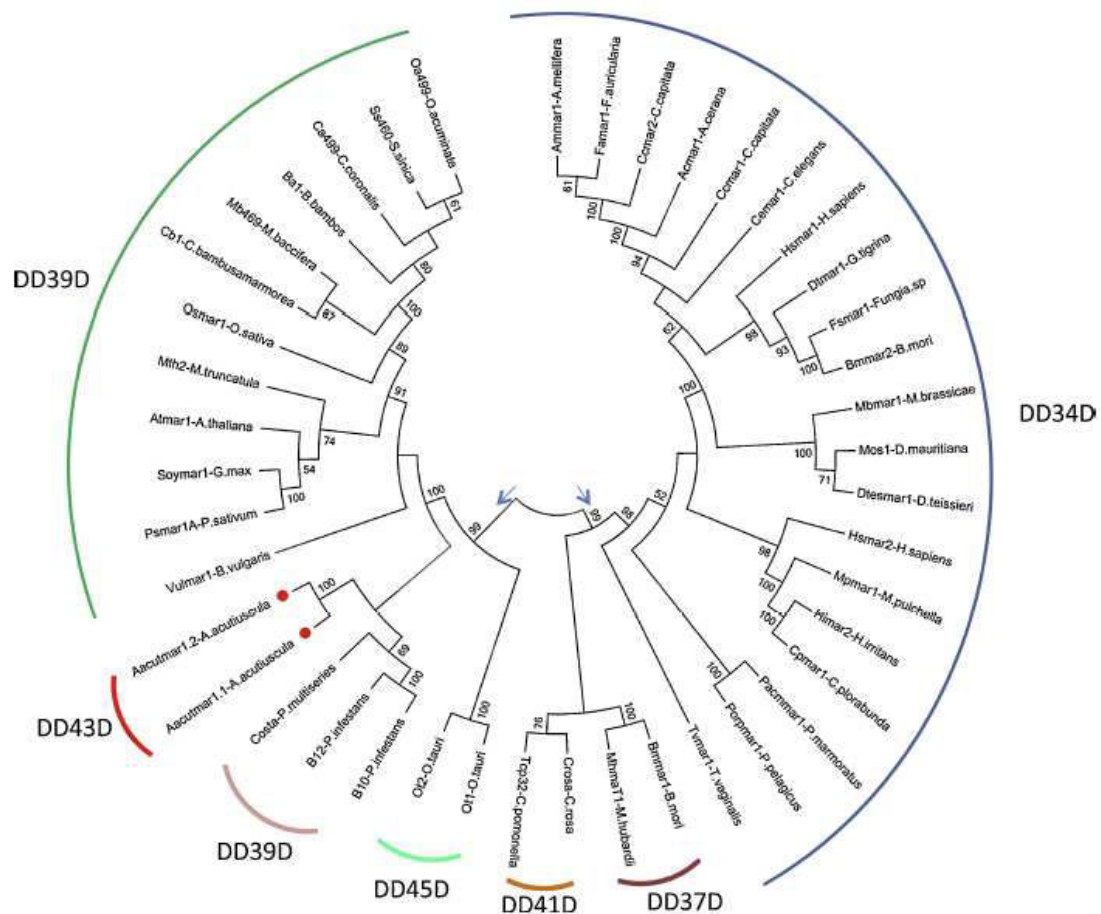
### Discussion

This study identified the first evidence of the presence of a complete *mariner*-like transposon (*Aacutmar1*) within the genome of the marine diatom *A. acutiuscula*. MLEs are widespread in eukaryotic genomes and partial fragments had already been detected in various microalgal species (Hermann 2011). *Costa*, a unique full-length MLE, was identified in silico in *Pseudonitzschia multiseri*. The MLEs identified in *A. acutiuscula* (2,114 bp) are shorter than similar transposons described in plants, such as *Soymar1* in soybean, with a length of 3,491 bp (Jarvik and Lark 1998), *Osmar1* in rice with 5,259 bp (Tarchini et al. 2000), *Vulmar1* in beet with 3,909 bp (Jacobs et al. 2004), *Psmar1* in pea with 3,204 bp (Macas

ConsAacutmar1	: MRA MPLFSAG-----TLHTIKM---	19
Costa	: MEAEKKHQHSTINLV-----QSIFLFLTIQARNN	31
Osmar1	: MPFFRTPL--SRMLLSPSIRIRSPSIQRTSPSATTHGICQGREQRTMDIDNLNLEPGDTYAATDPGEAGFDHSAATDAQNN	79
Phmar1	: MADPT-----DSGFDLNVREEDDD	20
Psmar1A	: M-----	1
Soymar1	: MQ-----	2
Vulmar1	: MDAAQEMY-----NAVDEEDL	17
ConsAacutmar1	: ----DQE-----NLEPNAKR-----VKTK	34
Costa	: NT-NNEQQ--TTHNN-----NNNSIMTSVAPPSEVLPPTTHEVLP-----RHQ	74
Osmar1	: GEAGFDHSA--TDAQNGEAGDFPFD--SNADDVVT-----EEIGVDPGQENVQPFDPFDDSDVDDAILVPQENTHKNV	154
Phmar1	: GNLFFDLNEPILEDHNNIDLNPLDEFGAVDFDYQNI-----AEQDVE-----APVQVHPKKH	75
Psmar1A	: -----	-
Soymar1	: -----	6
Vulmar1	: DEY-----YQLL-----E-----EIVPNFNSQISSEPE--SAAQDNKIKP	53
ConsAacutmar1	: ELTDNERIAVSMILIGM--SAKEPDEHGSFTDVAKKEGVRP--CVGELWAKATTSRFNGKVDEDEIKS--KKHERGRKMLWDI	112
Costa	: NITLNERQAVVHHLTPKQSTTSARTEKGAFFKESAALEGVTARTCSRVMWYLEVDPETCPAGVST--NRKNGGRKSTFA	153
Osmar1	: VLDNDKRAHEDAMTVK--ARKCYTKGHESKEVSASGSPVIRTVCRIWKKGKSCLDQG--ISV--TVASG--SRSGRKKVKV--	228
Phmar1	: DYFEHVRLVYQALLMR--SKNGRGNHNDTTIVSSQGGVKIRSVORINWQGNQLAQN--IPV--TVANLKKGRSGRKKATP--	150
Psmar1A	: -----YVYHLLQK--EVDGHRKGGATNEVASSNSVPLRTVORINWRAKESETH-----VVSRRKTKNGRKKTS--	64
Soymar1	: MGSNEERITVYQLLQK--EVDGHRKGGATNEVASSNSVCRKTDEIRWRAKESETH-----VVSRRKTKNGRKKTS--	77
Vulmar1	: EMKHEKRAKAVDELIM--EENGTEPYZALKKTAEEKVSRHTEIRWRAKESETH-----VVSRRKTKNGRKKTS--	128
<b>GRK motif</b>		
ConsAacutmar1	: DAMDEAVQQVRSVKKSYHGIAACGVPHITLIRYRE--ILVSHSNALKHTLIP--EHRAPRIEFMSHIDPNTVRNGSRKKY	192
Costa	: -AVEEAVKKSYKDRSDRTLSAASGLARTDGNVYKGCILRHSGAVKELFTEKNNKARVFPALSIIRK-----DGKF	226
Osmar1	: -VDVSCIEDHSLSRITQDVATQCVSTSRVVRMKKGAIKGVSSSDGHLTDNKRIDRLAWIDNLDERSVP--HNLVF	306
Phmar1	: -LDLEQANRHPKQV--IEDVSSRUGLSKSRITRYLKKGLRHSSTIKRYLDANKRTIRLWIDNIEQGV--IV--DDKFF	227
Psmar1A	: -IDENQIREFFPSORTNIRSAFAIKTNPASVPRILKSGATRENSNAIKELKEENKISRLFFLSMLGEG--TP--HDMFF	140
Soymar1	: -IDLQAREIPLSORTVIRTLAVAKTNTSAMVRLTOSGAIKHSSNAIKELKEENKISRLFFLSMLGEG--TP--HDMFF	153
Vulmar1	: -YSEFFZQSVPENLRTSVRCFAGAKVSPSAHVRILKFKLRSHTSHHRSLETHRLKEMVWLSHILPAHPL--YKQZF	206
ConsAacutmar1	: LPMLESVVDEKWFEMKDNQVETSKNREPERRHVQKCYGRKMLCAQRPFRMVDG--RM--GRKGRKWFSGYVGGPQR	271
Costa	: EDMYDTEIDEKWFILKTTNQTYYICSGPEPESHCKSKRFTEKVMMLCAVAPRPHADRNAGTIGKRGWFFVVKPQR	307
Osmar1	: KPIEDTIDEKWFILKRTVRYVAAPTASRRIRTIQNNKFPKIMMLTALAPRPFDSNGNCIFGKRGCFEFTVYTAAR	387
Phmar1	: RLLEDTIDEKWFILKQKSEYIYLPDEDEHRTCKNKYIFRMMLCVCAAPRFRN--GECVSGKRGCFEFTVYTAAR	307
Psmar1A	: KSMHNLIDEKWFILKCKSEYIYLPDEDEYRTCKNFKAKVMELVAQTRPFDSEENETSGKRGCFEFTVYTAAR	221
Soymar1	: QSMHNLIDEKWFILKCKSEYIYLPDEDEYRTCKNFKAKVMELVAQTRPFDSEENETSGKRGCFEFTVYTAAR	234
Vulmar1	: ILMENLIDEKWFILNPDTRFYLLRDEDEYRCCCKSKFKIKAMVIGVSKMYDAQNLHGRKGFEEFERVAKY	287
<b>D</b>		
ConsAacutmar1	: RSKHRPAGAVVRN--EYVKQVYRYVLELVIPATIDKFMRYKKEGVIIQDDGSTAHEDPGDEEENAMDLKRSRINI	351
Costa	: NSGNRPAGTMYVRN--EYVNAEYCKMMLVLPATIQEFPVCKWSREITIQEDNAKHHLRSRAVREAA--EAGVNI	384
Osmar1	: ESNRPAGTTEKMPESITKEVIRSFIRVLPVAVRAWPRDEAGKILIIQDDNARHHPADCPDCEAA--KDDGNI	464
Phmar1	: GSNRLAGECVIRPFCINRVIDDEMINVLPATIRAKWPHEDVHKIITIQDDNVPSHKVDPCQEPV--KDDGDI	384
Psmar1A	: SSNVRAGTMYRAITVNRDVPSFLIDVLPATIRKWPDEFEETIIQDDNARTHNNHDPLEEA--TKDGD	298
Soymar1	: SSNVRAGTMYRAITVNRDVPSFLIDVLPATIRKWPDEFEETIIQDDNARTHNNHDPLEEA--TKDGD	311
Vulmar1	: ESNRNGTTEERVVIRITQVIRDMIQNVLESMSEWBAQ--MSKTIQDDNARTHNOGNDEARAA--TCNEENI	363
<b>D</b>		
ConsAacutmar1	: ELVTOPCSPDNINIDLFERSIQILVYEAETEFALIKVDDAWYVPMATVIRWLYYSGINMIGHGEGNFSTEE	432
Costa	: EVVTOPSPDNINIDLFERSIQILVYKRAKNIDGLANVDEALVETSTIATNSIFSLICALEESRIVCGNNSKIRE	465
Osmar1	: ELVQPPNSPDNINIDLFERSIQILVYKSGATTEELVANIDRAEDFPVRLSRIFLS--EAMRNRIVVIGNSDLP	545
Phmar1	: ELVQPPNSPDNINIDLFERSIQILVYKDAKTILKELIPVQCALEGSPWKAIRFVTLCTVIRKIAMIRGCKNKKIPE	465
Psmar1A	: ELVQPPNSPDNINIDLFERSIQILVYKAPKTIDELISVVKSENGSPSIXSRIFVSLCLQIM--EVRKNGKNKKIPE	379
Soymar1	: ELVQPPNSPDNINIDLFERSIQILVYKAPKTIDELISVVKSENGSPSIXSRIFVSLCLQIM--EVRKNGKNKKIPE	392
Vulmar1	: ELVQPPNSPDNINIDLFERSIQILVYQTFEKNIDELITNVQDAQRDPSVMKYTWLCLQIM--EVLQVGGNNNNPE	444
<b>D</b>		
ConsAacutmar1	: LKREKLEKTLKLVLEVTEVAPSL-ELIQMQTEAGRRRSNRISGNSGDE-----EEEQQQNG	491
Costa	: LKREKLEKTLKLVLEVTEVADTVTFPQQHLAA-----LNMREV	503
Osmar1	: LKRGVLERQGLPLCLRCDAKSVNNANNV-----R	576
Phmar1	: LKRGVLERQGLPLCLRCDAKSVNNANNV-----AAN	499
Psmar1A	: LKREKLEKTLKLVLEVTEVAPSL-ELIQMQTEAGRRRSNRISGNSGDE-----ME	412
Soymar1	: LKREKLEKTLKLVLEVTEVAPSL-ELIQMQTEAGRRRSNRISGNSGDE-----NN	425
Vulmar1	: LKREKLEKTLKLVLEVTEVAPSL-ELIQMQTEAGRRRSNRISGNSGDE-----LPNESFQLEYFDNEGSLMKIKMNVCHVFVKITIFWKYIKSG	519

**Figure 5.** Alignment of the consensus amino acid sequence of the *ConsAacutmar* transposase and data bank sequences highlighting the conserved motifs: the GRK motif and the three motifs of the catalytic core; each D is indicated below the box. Costa JX013907; Osmar1 AF172282; Phmar1 HM581665; Psmar1 AY8333549; Soymar1 AF078934; Vulmar1 AJ556159.





**Figure 6.** Phylogenetic relationships of full-length *MLE* transposases; data collected in banks and isolated from protists, plants, animals and human. The tree was generated using UPGMA method, bootstrap values superior to 50% are indicated. The characteristic of the catalytic triad is indicated in the form “DDxD”, where x indicates the number of amino acid residues separating the two last aspartic acids. The arrows show the two main clades. The points identify *Aacutmar1* sequences.

et al. 2005), *Phmar1* in bamboo with 3,452 bp (Zhou et al. 2011) or recently *Costa* in the diatom *P. multiseries* with 3,035 bp (accession number JX013907).

*Aacutmar1* TIRs are 27 bp in length, and each is flanked by the characteristic *MLE* TSD, the dinucleotide TA. TIRs are imperfect, since four differences between the 5' and 3' regions can be observed. Imperfect TIRs are frequently observed among *MLEs* (Brillet et al. 2007; Plasterk et al. 1999), even in active sequences. For instance, the most famous active *MLE mos-1*, isolated from *Drosophila mauritiana* by Jacobson et al. (1986), exhibits 28-bp long imperfect TIRs with four differences. Comparison of *Aacutmar1* TIRs with

those of the six full-length sequences highlighted in plants and a diatom (*Soymar1*, *Osmar1*, *Vulmar1*, *Psmar1*, *Phmar1*, *Costa*) shows that the first ten nucleotides (5'-CTcCCTcYgt-3') are highly conserved even between very distant species, such as diatoms and rice. Feschotte et al. (2005) have shown that mutations located among the first nine nucleotides of *Osmar1* do not affect the binding of the transposase to TIRs. Thus, this short sequence does not seem to be involved in the interactions between a TIR and the enzyme.

Using TIRs as primers, 12 full-length elements were amplified in *A. acutiuscula* and two kinds of truncated sequences were also detected. These sequences (*Aacutmar2* and 3) are respectively

1,753 and 1,006 bp in length. In *Aacutmar2*, a large portion of the sequence has been replaced by a fragment that cannot be identified in data banks. In *Aacutmar3*, a large part of the transposase gene has been deleted. Even though *Aacutmar2* and 3 are deleted transposons, their intact TIRs allow them to be amplified using primers designed from TIRs, which suggests that they could be mobilized by a putative transposase produced by *Aacutmar1*.

Using quantitative PCR, the number of *Aacutmar* copies in the genome of *A. acutuscula* was estimated to be  $21 \pm 8$  copies for the full-length elements and  $38 \pm 14$  copies when including the deleted elements. This number of MLE copies is similar to that found in other species. In soybean, 25 copies of *Soymar1* were estimated by Southern-blot (Jarvik and Lark 1998), in *O. sativa* 34 sequences including 22 full-length elements have been identified (Feschotte et al. 2003) and in the bamboo *P. pubescens*, Zhou et al. (2011) detected the presence of 15 MLE copies.

After the *in silico* translation of the 12 full-length *Aacutmar1* sequences into amino acids, six elements showing an intact ORF were obtained. The putative transposase designated *AaTnp*, is 491 amino acids in length. In the DNA binding domain, two putative HTHs and a putative NLS have been revealed by *in silico* analysis at similar positions to those of other MLE transposases in data banks. The presence of two HTH motifs separated by a GRPR-like motif has been described in the *Tc1-mariner* superfamily (Plasterk et al. 1999; Richardson et al. 2006). This motif is replaced by the three residues GRK, which are highly conserved in *Aacutmar* and plant MLEs, and GRP in animals (Robertson 1995). The presence of HTH motifs is necessary for transposition and, in particular, for binding of the transposase onto TIRs, as has been demonstrated in *mos1* (Zhang et al. 2001).

In the catalytic domain of the transposase, the three canonical motifs surrounding the catalytic core are slightly different in *AaTnp* compared to MLEs from data banks, the differences are respectively: VDEKWF vs IDEKWF; IQQDGS vs IQ(Q/E)DN(A/V) and SPDTNINDLAFF vs SPD(L/F)N(V/I)DLGFF. Moreover, *AaTnp* exhibits a specific catalytic core of DD43D instead of the DD39D found in plant MLEs or in the diatom *P. multiseri*.

The description of most of the elements belonging to the *Tc1-mariner* superfamily revealed that the catalytic triad is DDx[D/E], where x ranges from 34 to 41 (Brillet et al. 2007). Two exceptions have been reported, one transposase with a DD45D motif in the green microalga *O. tauri* (Derelle et al.

2006) and, in the present study, evidence of a putative transposase with a DD43D motif in *A. acutuscula*. The phylogenetic relationship analysis showed that *Aacutmar1* clusters with the other elements isolated from Stramenopiles (*P. infestans* and *P. multiseri*). Hence the elements from *P. infestans* and *P. multiseri* do not cluster with the plant elements even if they share with them the same catalytic triad DD39D. A similar result was obtained by Coy and Tu (2005) who have shown that some elements sharing a DD43E triad belong to two clearly distinct families. Besides, the other MLE (*Costa*) identified in the diatom *P. multiseri* seems to be more closely related to the elements in the oomycete *P. infestans* than to *Aacutmar1*. This result, suggesting that diatoms may contain different MLEs, is however weakly supported (Fig. 6). The characterization of other MLEs in different diatom species is necessary to complete our knowledge about MLEs in diatoms and their relationships with the DNA transposons already identified.

## Conclusion

The present study describes new transposons belonging to the *Tc1-mariner* superfamily, designated *Aacutmar1*, 2, and 3, which were experimentally identified in the marine diatom *A. acutuscula*. Our analyses revealed: i) a significant proportion of full-length *Aacutmar1* among the elements detected; ii) several full-length elements with an intact ORF that suggests that they could still be active; iii) the presence of a TSD at each end of the copies; iv) a putative transposase containing two HTHs, a NLS, and a catalytic core with a catalytic triad DD43D, which differs from that of all other transposases described to date. *Aacutmar1* might be the first member of a new subfamily of the *Tc1-mariner* transposons.

## Methods

**Microalga and culture conditions:** *Amphora acutuscula* Kützinger (NCC216) originating from the southeast coast of Vietnam (Nguyen-Deroche et al. 2009, 2012) was grown in 500-mL Erlenmeyer flasks containing 250 mL of artificial sea water (ASW) (Harrison et al. 1980) at  $24 \pm 1^\circ\text{C}$  and an irradiance of  $300 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  (Li-Cor quantum meter Li-189) supplied by cool-white fluorescent tubes (Philips TLD 36 W) in a 14-10 h light-dark cycle. The calcium concentration was reduced from 9.4 mM to 0.25 mM in order to avoid adhesion of the cells to the flasks (Cooksey 1981). Axenic cultures were maintained in the exponential growth phase by diluting with fresh medium every 4 days before the beginning of experiments. The initial cell concentration was  $10^4 \text{ cells mL}^{-1}$ . After 4 days of culture, before the end of the exponential growth phase



(Nguyen-Deroche et al. 2009), cultures were sampled, frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until DNA extraction. The culture is available from the authors upon request.

**Genomic DNA extraction:** Cultures were centrifuged at 900 g with slow acceleration and deceleration that preserve cell integrity. Genomic DNA extractions were performed according to the protocol of Doyle and Doyle (1990). The gDNA pellets obtained were dried at room temperature and resuspended in Tris-EDTA buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The quality of nucleic acid extracts was assessed by an electrophoretic migration in 1% agarose gel. The gDNA purity and concentration were evaluated by spectrophotometry at 260 nm and 280 nm.

**PCR amplifications:** The PCR primer first used to search for full-length MLEs in *A. acutiuscula* was designed on the TIRs of Costa, a MLE transposon from *P. multiseriatus* (JX013907). The PCR primers used to search for partial *mariner* transposase genes in *A. acutiuscula* were those previously defined by Feschotte and Wessler (2002) to detect MLEs in flowering plants: MLE5A and MLE3A (Table 1). These degenerate primers are derived from conserved amino acid regions of the plant *mariner* transposases, IDEKWF and IQQDNA respectively. The full-length sequences were then amplified by PCR with primers defined in the TIR sequences (Fig. 1): a specific pair of primers (AcutTIR-5' and AcutTIR-3') or a single degenerate primer Acut-TIR). PCRs were performed in 50  $\mu\text{L}$ -reaction medium with 0.2  $\mu\text{M}$  of each primer, 100 ng of gDNA, and 2.5 U of *Taq* polymerase (Promega, Madison, WI). The PCR program consisted of an initial 5-min denaturing step at  $94^{\circ}\text{C}$ , followed by 40 cycles at  $94^{\circ}\text{C}$  each lasting 1 min, a 1 min step at the annealing temperature

(Table 1), 30 sec at  $72^{\circ}\text{C}$  for elongation process, and ending with 10 min at  $72^{\circ}\text{C}$ . PCR products were visualized on a 1% agarose gel with 0.5% ethidium bromide solution added before electrophoresis. PCR fragments were eluted from the gel using WIZARD SV Gel<sup>®</sup> and PCR Clean Up System<sup>®</sup> (Promega, Madison WI). The DNA fragments were then cloned in pGEM-T Vector System I<sup>®</sup> plasmid (Promega, Madison WI), and sequenced by the Cogenics-Genome express company (England, <http://www.cogenics.com>).

**Inverse PCR:** iPCR was used to elongate the partial MLE sequences previously amplified until full-length MLE sequences were obtained. 500 ng of *A. acutiuscula* gDNA were digested with 5 to 10 U of a restriction enzyme (EcoRI, Apal or EcoRV), and the resulting fragments were circularized in 150  $\mu\text{L}$  of a ligation buffer containing 100 U of T4-DNA ligase (Promega) at  $37^{\circ}\text{C}$  for 5 h. The enzyme was then inactivated at  $65^{\circ}\text{C}$  for 10 min, and 2  $\mu\text{L}$  of the ligated gDNA template were used for iPCR amplification (Ochman et al. 1988). *Mariner*-specific inverse nested primers were successively defined from the partial MLE sequences previously amplified and then used for the next round of iPCR (Table 1). The iPCR conditions were similar to those previously described: 3 min at  $94^{\circ}\text{C}$ , 35 cycles at  $94^{\circ}\text{C}$  for 30 s, the annealing temperature (Table 1) for 1 min, and  $72^{\circ}\text{C}$  for 2 min 30 s, followed by a final 10-min extension phase at  $72^{\circ}\text{C}$ . The iPCR products were sequenced and used to reconstruct the full-length MLE sequences by *in silico* assembly of the successive fragments obtained.

**Analysis of diatom MLE sequences:** The first full-length MLE sequence was reconstructed by successive assemblages of the sequences amplified by iPCR. The progressive assemblage process was carried out using pairwise alignment softwares available at (<http://multalin.toulouse.inra.fr/multalin/> and <http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Nucleotide sequences were conceptually translated into amino acids using a bioinformatic web service (<http://www.bioinformatics.org/>

<http://www.bioinformatics.org/sms/index.html>) and their identity with plant transposases was assessed using blastx at the NCBI site (<http://blast.ncbi.nlm.nih.gov>). MLE sequences were also compared with sequences available at the specific Transposable Elements databank Repbase (<http://www.girinst.org/censor/index.php>) and with the genomic database of *Phaeodactylum tricornutum* (strain CCMP632 / CCAP1055) (<http://genome.jgi-psf.org/Phatr2/Phatr2.home.html>). Putative HTH and NLS were searched using respectively GYM2.0 (Narasimhan et al. 2002; <http://www.cis.fiu.edu/~giri/bioinf/GYM2/prog.html>) and SeqNLS (Lin and Hu 2013; <http://mleg.cse.sc.edu/seqNLS>).

To detect complete transposons, TIRs were sought in the reconstructed sequences using a homemade algorithm developed by JD Rouault and based on the location of potential TSDs. To do this, all occurrences of TA located in an area between 0 and 1000 bp upstream and downstream of the transposase gene were searched for. Thus, all 20 nucleotide-sequences located just behind a first TA were aligned with the 20 nucleotides lying immediately in front of a second TA. For each pair of sequences identified, the number of differences was counted. If this number was less than a predetermined fixed threshold, e.g. 0, 2, 4, or 6, both TAs and their associated sequences were extracted and treated as potential TSDs and TIRs. When several pairs of identical or very close shots framing almost identical ORFs were found, the TIRs were precisely extended and aligned. Then, a consensual TIR was defined, and the differences from this consensus were identified. This statistical approach allowed us to define *in silico* TIRs, which could then be used to design primers to amplify the MLEs by PCR.

A genealogy of the full-length nucleotide sequences of *Aacutmar* was inferred using an algorithm that builds minimum spanning trees ([www.fluxus-engineering.com](http://www.fluxus-engineering.com)). As the set of sequences was a multistate data, the median-joining method was used to construct a phylogenetic network to illustrate the relationships between sequences (Bandelt et al. 1999). The distance between 2 sequences was measured by counting the number of character differences (Hamming distances).

An unrooted phylogenetic tree was created using the conceptual full-length amino-acid sequences of *A. acutiuscula* MLEs and *Tc1-mariner* elements already identified in different organisms. For this analysis, the *A. acutiuscula* MLE sequences *Aacutmar1.1* and *1.2* were compared with sequences from the infra-groups of the *Tc1-mariner* superfamily. The dataset included 43 sequences from the infra-groups identified according to the structure of the transposase gene as DDxD for MLEs referred as DD34D, DD37D, DD39D, DD41D, DD45D. Within the DD34D group, which is the most extensively studied, 2 or 3 elements were chosen from the main subfamilies *Irritans*, *Mauritiana*, *Mellifera*, *Elegans*, *Cecropia*, *Marmoratus*. For the other groups, at least 2 elements were chosen to complete the dataset when it was possible, except for the DD39D group in which all the known plant MLEs were added to the dataset since they provide the closest sequences to those of the Stramenopile MLEs.

All the sequences translated into amino acids for the analysis were aligned using Dambe 5 (<http://dambe.software.informer.com/5.2/>) and the unrooted phylogenetic tree was performed in MEGA version 5 based on UPGMA method (Tamura et al. 2011). Confidence of the groupings was estimated using 500 bootstrap replications. Only bootstrap values superior to 50% are shown (Fig. 6).

**Copy number estimation using real-time-PCR:** Before gDNA extraction, the cell density of each culture was determined using a Neubauer hemacytometer. Genomic DNA of four different cultures of *A. acutiuscula* was extracted from

7 µg of fresh microalgal cells as described by Doyle and Doyle (1990). Genomic DNA was dissolved in 350 µL of sterile water and treated by RNase following the manufacturer's instructions. The number of MLE copies in the genome of *A. acutiuscula* was determined using the  $\Delta C_t$  method (Eagle and Crease 2012) by comparing the rate of amplification of MLEs in the gDNA to that of the same amplicon inserted into the pGEM-T vector (calibrator plasmid), as described by Caprioara-Buda et al. (2012). The PCR target consisted of a specific fragment of the MLE open reading frame amplified using the primers Acutmar either F1/R1 or F2/R2 primers (Table 1). F1/R1 primers were used to quantify only *Aacutmar1* and the F2/R2 primers to quantify all the *Aacutmar* sequences. Plasmid suspensions with the PCR target were purified and quantified according to the method of Whelan et al. (2003), and sequenced (Cogenics, Beckman Coulter Genomic). Q-PCR was performed using the GoTaq QPCR Master Mix kit (Promega). 7.55 µL of Master Mix (SYBR Green, dNTPs, MgCl<sub>2</sub>, and GoTaq polymerase enzyme) were added to 0.8 µM of each primer and 0.15 µL of CXX (carboxy-X-rhodamine) to compensate for changes in fluorescence between samples according to the Promega protocol. 4.9 µL of template DNA (either gDNA or plasmid DNA) were added to the mix. Reactions were run on the StepOnePlus™ Real-Time PCR System (Applied Biosystems). The PCR program was as follows: 2 min at 95 °C, 45 15-s denaturing cycles at 95 °C, and 1 min annealing at 60 °C were performed. A dissociation curve was created during an additional cycle by increasing the temperature from 60 °C to 95 °C with increments of 0.3 °C. The threshold cycle (Ct) was determined by the Step One plus software version 2.1. Standard curves (dilution series of gDNA and plasmids) were generated to validate each primer pair and the percentile amplification efficiency (PAE) determined as follows:  $PAE = \ln(10^{(-1/\text{slope})})/\ln(2)$ . The  $\Delta C_t$  quantitative qPCR method (Eagle and Crease 2012) corrected by the PAE coefficient of Yuan et al. (2008) was used to estimate the number of MLE copies as  $2^{-\Delta C_t}$ , where  $\Delta C_t$  is  $((Ct \times PAE_{MLE \text{ in gDNA}}) - (Ct \times PAE_{MLE \text{ in plasmid}}))$ . The result is expressed as the mean of four independent experiments, and is calculated relative to the number of microalgal cells.

## Acknowledgements

This work was supported by the European FP7-KBBE-2010-4 GIAVAP contract and a grant of Vietnamese fellowship. The authors thank Monika Ghosh for correcting the English text and Wafa Ben-Lazhar for helpful discussions.

## References

- Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306:79–86
- Baldauf SL (2008) An overview of the phylogeny and diversity of eukaryotes. *J Syst Evol* 46:263–273
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. *Curr Opin Genet Dev* 15:621–627
- Biémont C, Vieira C (2006) Genetics: Junk DNA as an evolutionary force. *Nature* 443:521–524
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret J-P, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kröger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jézéquel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq M-P, Napoli C, Obornik M, Parker MS, Petit J-L, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siat M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van de Peer Y, Grigoriev IV (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456:239–244
- Böhne A, Brunet F, Galiana-Arnoux D, Schultheis C, Volff JN (2008) Transposable elements as drivers of genomic and biological diversity in vertebrates. *Chromosome Res* 16:203–215
- Brillet B, Bigot Y, Augé-Gouillou C (2007) Assembly of the *Tc1* and *mariner* transposition initiation complexes depends on the origins of their transposase DNA binding domains. *Genetica* 130:105–120
- Burki F, Shalchian-Tabrizi K, Pawlowski J (2008) Phylogenomics reveals a new «megagroup» including most photosynthetic eukaryotes. *Biol Lett* 4:366–369
- Caprioara-Buda M, Meyer W, Jeynov B, Corbisier P, Trapmann S, Emons H (2012) Evaluation of plasmid and genomic DNA calibrants used for the quantification of genetically modified organisms. *Anal Bioanal Chem* 404:29–42
- Capy P (2004) Structure et évolution des éléments transposables. *J Soc Biol* 198:393–398
- Capy P, Gasperi G, Biémont C, Bazin C (2000) Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85:101–106
- Carr M, Nelson M, Leadbeater BSC, Baldauf S (2008) Three family of LTR retrotransposons are present in the genome of the choanoflagellate *Monosiga brevicollis*. *Protist* 159:579–590
- Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot Microbiol* 46:347–366
- Cooksey KE (1981) Requirement for calcium in adhesion of a fouling diatom to glass. *Appl Environ Microbiol* 41:1378–1382



- Coy MR, Tu Z (2005) Gambol and *Tc1* are two distinct families of DD34E transposons: analysis of the *Anopheles gambiae* genome expands the diversity of the *IS630-Tc1-mariner* superfamily. *Insect Mol. Biol* 14:537–546
- Daboussi MJ, Capy P (2003) Transposable elements in filamentous fungi. *Annu Rev Microbiol* 57:275–299
- de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD (2011) Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 7:e1002384
- Deininger PL, Roy-Engel AM (2002) Mobile Elements in Animal and Plant Genomes. In: Craig NL, Craigie R, Gellert M, Lambowitz AM (eds) *Mobile DNA II*. ASM Press, Washington, DC, pp 1074–1091
- Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, Partensky F, Degroove S, Echeynie S, Cooke R, Saey Y, Wuyts J, Jabbari K, Bowler C, Panaud O, Piegu B, Ball S, Ral JP, Bouget FY, Piganeau G, De Baets B, Picard A, Delseny M, Demaille J, Van de Peer Y, Moreau H (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci USA* 103:11647–11652
- Deschamps P, Moreira D (2012) Reevaluating the green contribution to diatom genomes. *Genome Biol Evol* 4:683–688
- Doak TG, Doerder FP, Jahn CL, Herrick G (1994) A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common “D35E” motif. *Proc Natl Acad Sci USA* 91:942–946
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Eagle SHC, Crease TJ (2012) Copy number variation of ribosomal DNA and Pokey transposons in natural populations of *Daphnia*. *Mobile DNA* 3:4
- Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) The evolution of modern eukaryotic phytoplankton. *Science* 305:354–360
- Feschotte C, Pritham EJ (2005) Non-mammalian c-integrases are encoded by giant transposable elements. *Trends in Genetics* 21:551–552
- Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* 41:331–368
- Feschotte C, Wessler SR (2002) *Mariner*-like transposases are widespread and diverse in flowering plants. *Proc Natl Acad Sci USA* 99:280–285
- Feschotte C, Swamy L, Wessler SR (2003) Genome-wide analysis of *mariner*-like transposable elements in rice reveals complex relationships with stowaway miniature inverted repeat transposable elements (MITES). *Genetics* 163:747–758
- Feschotte C, Osterlund MT, Peeler R, Wessler SR (2005) DNA-binding specificity of rice *mariner*-like transposases and interactions with Stowaway MITES. *Nucleic Acids Res* 33:2153–2165
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281:237–240
- Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet* 5:103–107
- Finnegan DJ (1992) Transposable elements. *Curr Opin Genet Dev* 2:861–867
- Grandbastien MA, Audéon C, Bonnivard E, Casacuberta JM, Chalhoub B, Costa AP, Le QH, Melayah D, Petit M, Poncet C, Tam SM, Van Sluys MA, Mhiri C (2005) Stress activation and genomic impact of *Tnt1* retrotransposons in Solanaceae. *Cytogenet Genome Res* 110:229–241
- Granum E, Raven JA, Leegood RC (2005) How do marine diatom fix 10 billion tonnes of inorganic carbon per year? *Can J Bot* 83:898–908
- Goodwin TJ, Butler MI, Poulter RT (2003) *Cryptons*: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. *Microbiology* 149:3099–3109
- Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, Bhattacharya D (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with chromalveolates. *Mol Biol Evol* 24:1702–1713
- Harrison PJ, Waters RE, Taylor FJR (1980) A broad spectrum artificial seawater medium for coastal and open ocean phytoplankton. *J Phycol* 16:28–35
- Henikoff S, Plasterk RH (1988) Related transposons in *C. elegans* and *D. melanogaster*. *Nucleic Acids Res* 16, 6234–6234
- Hermann D (2011) Caractérisation d'éléments transposables de type *mariner* chez les microalgues marines. PhD thesis, Le Mans University (France), 268p
- Hua-Van A, Davière JM, Kaper F, Langin T, Daboussi MJ (2000) Genome organization in *Fusarium oxysporum*: clusters of class II transposons. *Curr Genet* 37:339–347
- Ivics Z, Li MA, Mátés L, Boeke JD, Nagy A, Bradley A, Izsvák Z (2009) Transposon-mediated genome manipulation in vertebrates. *Nat Methods* 6:415–422
- Jacobs G, Dechyeva D, Menzel G, Dombrowski C, Schmidt T (2004) Molecular characterization of *Vulmar1*, a complete *mariner* transposon of sugar beet and diversity of *mariner*- and *En/Spm*-like sequences in the genus *Beta*. *Genome* 47:1192–1201
- Jacobson JW, Medhora MM, Hartl DL (1986) Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc Natl Acad Sci USA* 83:8684–8688
- Jarvik T, Lark KG (1998) Characterization of *Soymar1*, a *mariner* element in soybean. *Genetics* 149:1569–1574
- Kapitonov VV, Jurka J (2001) Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci USA* 98:8714–8719
- Kapitonov VV, Jurka J (2006) Self-synthesizing DNA transposons in eukaryotes. *Proc Natl Acad Sci USA* 103:4540–4545
- Kapitonov VV, Jurka J (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet* 9:411–412
- Keeling PJ, Archibald JM, Fast NM, Palmer JD (2004) Comment on “The evolution of modern eukaryotic phytoplankton”. *Science* 306:2191
- Kentner EK, Arnold ML, Wessler SR (2003) Characterization of high-copy-number retrotransposons from the large genomes of the *Louisiana iris* species and their use as molecular markers. *Genetics* 164:685–697

- Kidwell MG (2002) Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115:49–63
- Kidwell MG, Lisch D (1997) Transposable elements as sources of variation in animals and plants. *Proc Natl Acad Sci USA* 94:7704–7711
- Kidwell MG, Lisch DR (2001) Perspectives: transposable elements, parasitic DNA, and genome. *Evolution* 55:1–24
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921
- Lane CE, Archibald JM (2008) The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol Evol* 23:268–275
- Laudencia-Chingcuanco D, Fowler DB (2012) Genotype-dependent burst of transposable element expression in crowns of hexaploid wheat (*Triticum aestivum* L.) during cold acclimation. *Comp Funct Genom*, Article ID 232530
- Lin J, Hu J (2013) SeqNLS: Nuclear localization signal prediction based on frequent pattern mining and linear motif scoring. *PLoS ONE* 8(10):e76864
- Lönning WE, Saedler H (1997) Plant transposons: contributors to evolution? *Gene* 205:245–253
- Lönning WE, Saedler H (2002) Chromosome rearrangements and transposable elements. *Annu Rev Genet* 36:389–410
- Macas J, Koblízková A, Neumann P (2005) Characterization of Stowaway MITEs in pea (*Pisum sativum* L.) and identification of their potential master elements. *Genome* 48:831–839
- Maumus F, Allen AE, Mhiri C, Hu H, Jabbari K, Vardi A, Grandbastien MA, Bowler C (2009) Potential impact of stress activated retrotransposons on genome evolution in a marine diatom. *BMC Genomics* 10:624
- Narasimhan G, Bu C, Gao Y, Wang X, Xu N, Mathee K (2009) Mining for Motifs in Protein Sequences. *J Comput Biol* 9:707–720
- Nguyen-Deroche TLN, Le TT, Bui TV, Rincé Y, Tremblin G, Morant-Manceau A (2009) Effects of copper on growth and photosynthesis in marine diatoms: a comparison between species from two different geographical areas. *Cryptogam Algal* 30:97–109
- Nguyen-Deroche TLN, Caruso A, Le TT, Bui TV, Schoefs B, Tremblin G, Morant-Manceau A (2012) Zinc affects differently growth, photosynthesis, antioxidant enzyme activities and phytochelatin synthase expression of four marine diatoms. *The Sci World J* 2012, Article ID 982957
- Nisbet RER, Kilian O, McFadden GI (2004) Diatom genomics: genetic acquisitions and mergers. *Curr Biol* 14:R1048–R1050
- Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120:621–623
- Pace JK, Gilbert C, Clark MS, Feschotte C (2008) Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci USA* 105:17023–17028
- Parfrey LW, Barbero E, Lasser E, Dunthorn M, Bhattacharya D, Patterson DJ, Katz LA (2006) Evaluating support for the current classification of eukaryotic diversity. *PLoS Genet* 2:e220
- Plasterk RH, Izsvák Z, Ivics Z (1999) Resident aliens: the *Tc1-mariner* superfamily of transposable elements. *Trends Genet* 15:326–332
- Ray DA, Feschotte C, Pagan HJ, Smith JD, Pritham EJ, Arensburger P, Atkinson PW, Craig NL (2008) Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Res* 18:717–728
- Reyes-Prieto A, Weber AP, Bhattacharya D (2007) The origin and establishment of the plastid in algae and plants. *Annu Rev Genet* 41:147–168
- Richardson JM, Dawson A, O'Hagan N, Taylor P, Finnegan DJ, Walkinshaw MD (2006) Mechanism of *Mos1* transposition: insights from structural analysis. *EMBO J* 25:1324–1334
- Robertson HM (1995) The *Tc1-mariner* superfamily of transposons in animals. *J Insect Physiol* 41:99–105
- Robertson HM (2002) Evolution of DNA Transposons in Eukaryotes. In Craig NL, Craigie R, Gellert M, Lambowitz AM (eds) *Mobile DNA II*. ASM Press, Washington, DC, pp 1093–1110
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambrose C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Bruntnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddelloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112–1115
- Seidl MF, Van den Ackerveken G, Govers F, Snel B (2012) Reconstruction of oomycete genome evolution identifies differences in evolutionary trajectories leading to present-day large gene family. *Genome Biol Evol* 4:199–211
- Shao H, Tu Z (2001) Expanding the diversity of the *IS630-Tc1-mariner* superfamily: discovery of a unique DD37E transposon and reclassification of the DD37D and DD39D transposons. *Genetics* 159:1103–1115
- Silva JC, Bastida F, Bidwell SL, Johnson PJ, Carlton JM (2005) A potentially functional *mariner* transposable element in the protist *Trichomonas vaginalis*. *Mol Biol Evol* 22:126–134

- Simon N, Cras AL, Foulon E, Lemée R (2009) Diversity and evolution of marine phytoplankton. *C R Biol* 332: 159–170
- Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8:272–285
- Tamura K, Petreson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA 5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood evolutionary distance and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Tarchini R, Biddle P, Wineland R, Tingey S, Rafalski A (2000) The complete sequence of 340 kb of DNA around the rice *Adh1-adh2* region reveals interrupted colinearity with maize chromosome 4. *Plant Cell* 12: 381–391
- Tenaillon MJ, Hollister JD, Gaut BS (2010) A triptych of the evolution of plant transposable elements. *Trends Plant Sci* 15:471–478
- Tirichine L, Bowler C (2011) Decoding algal genomes: tracing back the history of photosynthetic life on Earth. *Plant J* 66: 45–57
- Venner S, Feschotte C, Biémont C (2009) Dynamics of transposable elements: towards a community ecology of the genome. *Trends Genet* 25:317–323
- Vitte C, Panaud O, Quesneville H (2007) LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss. *BMC Genomics* 8:218
- Whelan JA, Russell NB, Whelan MA (2003) A method for the absolute quantification of cDNA using real-time PCR. *J Immunol Methods* 278:261–269
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chal-houb B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8:973–982
- Yuan JS, Wang D, Stewart CN Jr (2008) Statistical methods for efficiency adjusted real-time PCR quantification. *Biotechnol J* 3:112–123
- Zhang L, Dawson A, Finnegan DJ (2001) DNA-binding activity and subunit interaction of the *mariner* transposase. *Nucleic Acids Res* 29:3566–3575
- Zhou MB, Zhong H, Tang DQ (2011) Isolation and characterization of seventy-nine full-length *mariner*-like transposase genes in the Bambusoideae subfamily. *J Plant Res* 124:607–617
- Zuccolo A, Sebastian A, Talag J, Yu Y, Kim H, Collura K, Kudrna D, Wing RA (2007) Transposable element distribution, abundance and role in genome size variation in the genus *Oryza*. *BMC Evol Biol* 7:152–167

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect



# Thèse de Doctorat

Duc Hung NGUYEN

## Caractérisation et expression de nouveaux éléments génétiques transposables de la superfamille *Tc1-mariner* chez la microalgue marine *Amphora acutiuscula* (Bacillariophyta)

### Résumé

Les éléments génétiques transposables (ET) sont des séquences d'ADN capables de se déplacer dans tous les génomes sous certaines conditions. Les ET ont des structures et des modes de transposition qui les différencient en plusieurs groupes. Les éléments de la famille *mariner* sont ubiquistes ; lorsqu'ils sont actifs, ils produisent une transposase qui coupe l'élément et l'insère dans un autre locus. Des fragments d'éléments de type *mariner* (MLE) ont été précédemment identifiés chez la diatomée marine *Amphora acutiuscula* et ils semblaient actifs en condition de stress thermiques. Les diatomées sont caractérisées par une paroi siliceuse ornementée et, dans le milieu marin, elles jouent un rôle primordial dans les chaînes alimentaires.

Dans le présent travail, par des méthodes de biologie moléculaire et de bioinformatique, nous avons recherché et caractérisé des MLE complets, et précisé leur activité en conditions de chocs thermiques et métalliques (cuivre et zinc). L'analyse des séquences amplifiées a mis en évidence la présence de MLE particuliers chez cette diatomée qui code une transposase ayant une triade catalytique DD43D jamais décrite jusqu'à présent. L'analyse phylogénétique place les MLE de diatomées dans une sousfamille différente mais proche de celle des MLE de plantes. Parmi les nombreuses copies de MLE présentes dans le génome d'*A. acutiuscula*, certaines sont exprimées lorsque la diatomée est placée pendant 2 à 5 heures à une température inférieure à sa température de culture. Par contre, l'expression des MLE n'a pas été mise en évidence chez cette espèce soumise aux stress métalliques appliqués.

**Mots clés :** élément transposable, *mariner*-like element, diatomée marine, stress thermique, stress métallique.

### Abstract

Transposable elements (TE) are DNA sequences able to move in all genomes depending on conditions. TE have different structures and transposition mechanisms. *Tc1-mariner* elements are ubiquitous; when they are active, they produce a transposase which cuts and inserts the element into another locus. Fragments of *mariner*-like elements (MLE) had previously been identified in the marine diatom *Amphora acutiuscula* and they seemed active under thermal stress. Diatoms are characterized by a siliceous ornamented cell wall and, in the ocean, they play a major role in trophic networks.

In this work, with biomolecular and bioinformatic methods, we have searched for and characterized full length MLE, and precised their activity under thermal and metal (copper and zinc) stresses. The analyse of the DNA sequences obtained highlighted that MLE in diatoms are particular and that they encode a transposase which has a DD43D catalytic triad never so far depicted. The phylogenetic analyse arranged diatom MLE in a subfamily different and close to that of plant MLE. Among the numerous MLE copies present in the genome of *A. acutiuscula*, some were expressed when the microalga was put at a lower temperature than the culture temperature for 2 to 5 hours. Conversely, metal stresses we applied did not induce MLE expression in this species.

**Key Words:** transposable element, *mariner*-like element, marine diatom, thermal stress, metal stress.